

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 15:27:32 ; Search time 28 seconds
(without alignments)
73.588 Million cell updates/sec

Title: US-09-991-433-5

Perfect score: 55

Sequence: 1 NKGTOQYTDQ 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Minimum number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SPTRMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_plant:*
11: sp_ricent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	138	12	085155 human parvo
2	55	100.0	141	12	085146 human parvo
3	55	100.0	141	12	085166 human parvo
4	55	100.0	141	12	085171 human parvo
5	55	100.0	142	12	085168 human parvo
6	55	100.0	142	12	085173 human parvo
7	55	100.0	144	12	085138 human parvo
8	55	100.0	145	12	085161 human parvo
9	55	100.0	145	12	085181 human parvo
10	55	100.0	146	12	085158 human parvo
11	55	100.0	147	12	085142 human parvo
12	55	100.0	148	12	085150 human parvo
13	55	100.0	151	12	085131 human parvo
14	55	100.0	151	12	085177 human parvo
15	55	100.0	153	12	085123 human parvo
16	55	100.0	157	12	085196 human parvo

17	55	100.0	162	12	085135 human parvo
18	55	100.0	546	12	0913X0 human parvo
19	55	100.0	546	12	0913W7 human parvo
20	55	100.0	554	12	0913P7 human parvo
21	55	100.0	554	12	09WKL9 human parvo
22	55	100.0	554	12	09PZS9 human parvo
23	55	100.0	554	12	0912B7 human eryth
24	55	100.0	554	12	065790 human parvo
25	55	100.0	554	12	090201 human parvo
26	55	100.0	760	12	09PZT8 human parvo
27	55	100.0	761	12	09PZU0 human parvo
28	55	100.0	765	12	09PZT6 human parvo
29	55	100.0	769	12	09PZT4 human parvo
30	55	100.0	773	12	0913X1 human parvo
31	55	100.0	773	12	0913W8 human parvo
32	55	100.0	781	12	09UGS0 human parvo
33	55	100.0	781	12	09JGP8 human parvo
34	55	100.0	781	12	09WKM0 human parvo
35	55	100.0	781	12	09PZT0 human parvo
36	55	100.0	781	12	0912B8 human eryth
37	55	100.0	781	12	085316 human parvo
38	55	100.0	781	12	085317 human parvo
39	55	100.0	781	12	085318 human parvo
40	55	100.0	781	12	085319 human parvo
41	55	100.0	781	12	085320 human parvo
42	55	100.0	781	12	085321 human parvo
43	55	100.0	781	12	090221 human parvo
44	55	100.0	781	12	090222 human parvo
45	55	100.0	781	12	090223 human parvo

ALIGNMENTS

RESULT 1
ID 085155 PRELIMINARY; PRT; 138 AA.
AC 085155;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Pobilczki A., Gigler A., Cassinotti P., Siegl G.,
RT "XXXSequence variability among different parvovirus B19 isolates."
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z70564; CA94477.1; -
DR InterPro: IPR001403; Parvo Coat.
DR Pfam: PF00740; Parvo Coat. 1.
FT NON_TER 1
FT NON_TER 138
FT NON_TER 138
SQ SEQUENCE 138 AA; 15237 MW; 5899FBB879A3E68D CRC64;
Query Match 100.0%; Score 55; DB 12; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.005; Mismatches 0; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NKGTOQYTDQ 10
Db 107 NKGTOQYTDQ 116
RESULT 2
ID 085146 PRELIMINARY; PRT; 141 AA.
AC 085146;
DT 01-NOV-1996 (TREMblrel. 01, Created)

DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses: ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 ON NCBI_TaxId=10798;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotzki A., Giegler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z70555; CA94467.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON_TER 1 141
 FT NON_TER 141 141
 SQ SEQUENCE 141 AA; 15712 MW; 32F6B23E7E6B8583 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 55; DB 12; Length 141;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTQD 10
 DB 106 NKGTQOYTQD 115

RESULT 3
 ID 085166 PRELIMINARY; PRT; 141 AA.
 AC 085166;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses: ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 ON NCBI_TaxId=10798;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotzki A., Giegler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z70575; CA94488.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON_TER 1 141
 FT NON_TER 141 141
 SQ SEQUENCE 141 AA; 15785 MW; 4115D3D915751757 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 55; DB 12; Length 141;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTQD 10
 DB 109 NKGTQOYTQD 118

RESULT 4
 ID 085171 PRELIMINARY; PRT; 141 AA.
 AC 085171;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.

OC Viruses: ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 ON NCBI_TaxId=10798;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotzki A., Giegler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z70580; CA94493.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON_TER 1 141
 FT NON_TER 141 141
 SQ SEQUENCE 141 AA; 15770 MW; C9B92572A78C6C27 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 55; DB 12; Length 141;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTQD 10
 DB 109 NKGTQOYTQD 118

RESULT 5
 ID 085168 PRELIMINARY; PRT; 142 AA.
 AC 085168;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses: ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 ON NCBI_TaxId=10798;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotzki A., Giegler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z70577; CA94490.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON_TER 1 142
 FT NON_TER 142 142
 SQ SEQUENCE 142 AA; 15821 MW; 2B6E4D9A6784F8C4 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 55; DB 12; Length 142;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTQD 10
 DB 105 NKGTQOYTQD 114

RESULT 6
 ID 085173 PRELIMINARY; PRT; 142 AA.
 AC 085173;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses: ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 ON NCBI_TaxId=10798;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotzki A., Giegler A., Cassinotti P., Siegl G.,

RA Wolf H., Modrow S.;
RT "xxxSequence variability among different parvovirus B19 isolates.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z70582; CA94495.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
FT NON_TER 1 1
FT NON_TER 142 142
SQ SEQUENCE 142 AA: 15820 MW: 3A9F5E07C0BC0434 CRC64;

Query Match 100.0%; Score 55; DB 12; Length 142;
Best Local Similarity 100.0%; Pred. No. 0.0052;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKGTOOYTDQ 10
DB 105 NKGTOOYTDQ 114

RESULT 7

ID 085138 PRELIMINARY; PRT; 144 AA.
AC 085138;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Poblitzki A., Gigler A., Cassinotti P., Siegl G.,
RA Wolf H., Modrow S.;
RT "xxxSequence variability among different parvovirus B19 isolates.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z70547; CA94459.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
FT NON_TER 1 1
SQ SEQUENCE 144 AA: 16022 MW: 95D6EF37BA5997AE CRC64;

Query Match 100.0%; Score 55; DB 12; Length 144;
Best Local Similarity 100.0%; Pred. No. 0.0053;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKGTOOYTDQ 10
DB 107 NKGTOOYTDQ 116

RESULT 8

ID 085161 PRELIMINARY; PRT; 145 AA.
AC 085161;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Poblitzki A., Gigler A., Cassinotti P., Siegl G.,
RA Wolf H., Modrow S.;
RT "xxxSequence variability among different parvovirus B19 isolates.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z70570; CA94483.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.

FT NON_TER 1 1
FT NON_TER 145 145
SQ SEQUENCE 145 AA: 16165 MW: B4FE95AB6EAD9A67 CRC64;

Query Match 100.0%; Score 55; DB 12; Length 145;
Best Local Similarity 100.0%; Pred. No. 0.0053;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKGTOOYTDQ 10
DB 105 NKGTOOYTDQ 114

RESULT 9

ID 085181 PRELIMINARY; PRT; 145 AA.
AC 085181;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Poblitzki A., Gigler A., Cassinotti P., Siegl G.,
RA Wolf H., Modrow S.;
RT "xxxSequence variability among different parvovirus B19 isolates.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z70590; CA94503.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
FT NON_TER 1 1
FT NON_TER 145 145
SQ SEQUENCE 145 AA: 16210 MW: 6B45A1E1B9E923C4 CRC64;

Query Match 100.0%; Score 55; DB 12; Length 145;
Best Local Similarity 100.0%; Pred. No. 0.0053;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKGTOOYTDQ 10
DB 109 NKGTOOYTDQ 118

RESULT 10

ID 085158 PRELIMINARY; PRT; 146 AA.
AC 085158;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Poblitzki A., Gigler A., Cassinotti P., Siegl G.,
RA Wolf H., Modrow S.;
RT "xxxSequence variability among different parvovirus B19 isolates.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z70567; CA94480.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
FT NON_TER 1 1
FT NON_TER 146 146
SQ SEQUENCE 146 AA: 16252 MW: 2B74FE95AB6EAD9A CRC64;

Query Match 100.0%; Score 55; DB 12; Length 146;

Best Local Similarity 100.0%; Pred. No. 0.0053;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDO 10
DB 105 NKGTQOYTDO 114

RESULT 11

085142

ID 085142; PRELIMINARY; PRT; 147 AA.

AC 085142; 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DR 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE Viral protein 1 (Fragment).

GN VP1.

OS Human parvovirus B19.

OC Viruses; SSDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

NCBI_TaxID=10798;

[1]

SEQUENCE FROM N.A.

RA Hemner A., Von Poblotzki A., Giegler A., Cassinotti P., Siegl G.,

Wolff H., Modrow S.; "XXSsequence variability among different parvovirus B19 isolates.";

RT Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.

RL EMBL; 270551; CAA94463.1; -.

DR InterPro: IPR001403; Parvo_coat.

DR Pfam: PF00740; Parvo_coat; 1.

FT NON_TER 1 1

FT NON_TER 147 147

SQ SEQUENCE 147 AA; 16424 MW; E4ECF2459B308BA9 CRC64;

QY 1 NKGTQOYTDO 10
DB 108 NKGTQOYTDO 117

RESULT 12

085150

ID 085150; PRELIMINARY; PRT; 148 AA.

AC 085150; 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DR 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE Viral protein 1 (Fragment).

GN VP1.

OS Human parvovirus B19.

OC Viruses; SSDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

NCBI_TaxID=10798;

[1]

SEQUENCE FROM N.A.

RA Hemner A., Von Poblotzki A., Giegler A., Cassinotti P., Siegl G.,

Wolff H., Modrow S.; "XXSsequence variability among different parvovirus B19 isolates.";

RT Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.

RL EMBL; 270559; CAA94471.1; -.

DR InterPro: IPR001403; Parvo_coat.

DR Pfam: PF00740; Parvo_coat; 1.

FT NON_TER 1 1

FT NON_TER 148 148

SQ SEQUENCE 148 AA; 16539 MW; 8814ECF2459B308B CRC64;

QY 1 NKGTQOYTDO 10
DB 109 NKGTQOYTDO 118

DB 108 NKGTQOYTDO 117

RESULT 13

085131

ID 085131; PRELIMINARY; PRT; 151 AA.

AC 085131; 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DR 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE Viral protein 1 (Fragment).

GN VP1.

OS Human parvovirus B19.

OC Viruses; SSDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

NCBI_TaxID=10798;

[1]

SEQUENCE FROM N.A.

RX MEDLINE=96332516; PubMed=8760426;

RA Hemner A., Von Poblotzki A., Giegler A., Cassinotti P., Siegl G.,

Wolff H., Modrow S.; "Sequence variability among different parvovirus B19 isolates.";

RT J. Gen. Virol. 77:1781-1785(1996).

RL EMBL; 270540; CAA94452.1; -.

DR InterPro: IPR001403; Parvo_coat.

DR Pfam: PF00740; Parvo_coat; 1.

FT NON_TER 1 1

FT NON_TER 151 151

SQ SEQUENCE 151 AA; 16902 MW; 3CA74914BBE73A3E CRC64;

QY 1 NKGTQOYTDO 10
DB 109 NKGTQOYTDO 118

RESULT 14

085177

ID 085177; PRELIMINARY; PRT; 151 AA.

AC 085177; 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DR 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE Viral protein 1 (Fragment).

GN VP1.

OS Human parvovirus B19.

OC Viruses; SSDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

NCBI_TaxID=10798;

[1]

SEQUENCE FROM N.A.

RA Hemner A., Von Poblotzki A., Giegler A., Cassinotti P., Siegl G.,

Wolff H., Modrow S.; "XXSsequence variability among different parvovirus B19 isolates.";

RT Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.

RL EMBL; 270586; CAA94499.1; -.

DR InterPro: IPR001403; Parvo_coat.

DR Pfam: PF00740; Parvo_coat; 1.

FT NON_TER 1 1

FT NON_TER 151 151

SQ SEQUENCE 151 AA; 16888 MW; 06C6A614A616DDA3 CRC64;

QY 1 NKGTQOYTDO 10
DB 109 NKGTQOYTDO 118

RESULT 15

```

Q85123
ID Q85123 PRELIMINARY; PRT; 153 AA.
AC Q85123:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96332516; PubMed=8760426;
RA Hemauer A., Von Pohlitzki A., Giegler A., Cassinotti P., Siegl G.,
  Wolf H., Modrow S.;
  "Sequence variability among different parvovirus B19 isolates.";
  J. Gen. Virol. 77:1781-1785(1996).
RL EMBL; 270532; CAA94444.1; -
DR InterPro: IPR001403; Parvo_coat.
PF Pfam: PF00740; Parvo_coat; 1.
FT NON_TER 1
FT NON_TER 153
SQ SEQUENCE 153 AA; 17131 MW; ECCA1F44020814EC CRC64;

Query Match 100.0%; Score 55; DB 12; Length 153;
Best Local Similarity 100.0%; Pred. No. 0.0056;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQ 10
   |||
Db 108 NKGTOQYTDQ 117

```

Search completed: December 11, 2002, 15:29:37
 Job time : 28 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 15:27:07 ; Search time 11 seconds

(without alignments)
37.706 Million cell updates/sec

Title: US-09-991-433-5
Perfect score: 55
Sequence: .1 NKGTQYTDQ 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	781	1	COAT_PAVHB
2	54	61.8	546	1	CH60_ACTAC
3	54	61.8	546	1	CH60_ACTAC
4	54	61.8	547	1	CH60_PASMU
5	54	61.8	548	1	CH60_PASMU
6	34	61.8	652	1	C3BB_BACTO
7	34	61.8	659	1	C3BA_BACTO
8	33	60.0	872	1	CLPE_SYNY3
9	33	60.0	224	1	GSNU_CAEEL
10	33	60.0	238	1	TRY3_SALSA
11	33	60.0	244	1	TRY2_XENLA
12	33	60.0	246	1	UBIG_XYLA
13	33	60.0	310	1	IFRH_TOBAC
14	33	60.0	424	1	CRU3_ARATH
15	33	60.0	445	1	6PGD_CITAM
16	33	60.0	445	1	6PGD_CITAM
17	33	60.0	445	1	6PGD_CITAM
18	33	60.0	445	1	6PGD_CITAM
19	33	60.0	445	1	6PGD_CITAM
20	33	60.0	445	1	6PGD_CITAM
21	33	60.0	445	1	6PGD_CITAM
22	33	60.0	445	1	6PGD_CITAM
23	33	60.0	445	1	6PGD_CITAM
24	33	60.0	445	1	6PGD_CITAM
25	33	60.0	445	1	6PGD_CITAM
26	33	60.0	445	1	6PGD_CITAM
27	33	60.0	445	1	6PGD_CITAM
28	33	60.0	445	1	6PGD_CITAM
29	33	60.0	445	1	6PGD_CITAM
30	33	60.0	445	1	6PGD_CITAM
31	33	60.0	445	1	6PGD_CITAM
32	33	60.0	445	1	6PGD_CITAM
33	33	60.0	445	1	6PGD_CITAM

34	32	58.2	290	1	MREC_BACSU	001466 bacillus su
35	32	58.2	487	1	YOMI_CAEEL	P30651 caenorhabd
36	32	58.2	1063	1	PDOM_CIAAP	P22373 aradiceps p
37	32	58.2	1164	1	PHVD_ARATH	P42437 arabidopsi
38	31	56.4	200	1	PHNH_RHIME	O52985 rhizobium m
39	31	56.4	204	1	YPOL_ACICA	P07778 acinetobact
40	31	56.4	288	1	SSRA_ONCMY	P45433 oncorhynch
41	31	56.4	330	1	PHLC_STRAU	P09978 staphylococ
42	31	56.4	360	1	HISB_BACSU	P17731 bacillus su
43	31	56.4	428	1	HDA3_MOUSE	O88895 mus musculu
44	31	56.4	428	1	HDA3_CHICK	P56520 gallus gall
45	31	56.4	428	1	HDA3_HUMAN	O15379 homo sapien

ALIGNMENTS

RESULT 1	COAT_PAVHB	STANDARD;	PRT;	781 AA.
ID	COAT_PAVHB			
AC	P07299;			
DT	01-APR-1988 (Rel. 07, Created)			
DT	01-APR-1988 (Rel. 07, last sequence update)			
DT	01-APR-1990 (Rel. 14, last annotation update)			
DE	Probable coat protein VPI.			
OS	Human parvovirus B19.			
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.			
OX	NCBI_TaxID=10798;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Isolate AU;			
RX	MEDLINE=86200451; PubMed=3701931;			
RA	Shade R.O., Blundell M.C., Cotmore S.F., Tattersall P., Astell C.R.;			
RT	"Nucleotide sequence and genome organization of human parvovirus B19			
RT	isolated from the serum of a child during aplastic crisis.";			
RL	J. Virol. 58:921-936(1986).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: M13178; AAA66867.1; -.			
DR	PIR: A24299; VCPV19.			
DR	InterPro: IPR001403; Parvo_coat.			
DR	Pfam: PF00740; Parvo_coat; 1.			
KW	Coat protein; Glycoprotein.			
FT	CARBOHYD 46			
FT	CARBOHYD 184			
FT	CARBOHYD 220			
FT	CARBOHYD 293			
FT	SEQUENCE 781 AA; 86015 MW; 8062545D0576B07 CRC64;			
QY	1 NKGTQYTDQ 10			
DB	622 NKGTQYTDQ 631			
Query Match	100.0%;	Score 55;	DB 1;	Length 781;
Best Local Similarity	100.0%;	Pred. No. 0.003;		
Matches 10;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
RESULT 2	CH60_ACTAC	STANDARD;	PRT;	546 AA.
ID	CH60_ACTAC			
AC	P46398;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-OCT-1996 (Rel. 34, last sequence update)			
DT	16-OCT-2001 (Rel. 40, last annotation update)			
DE	60 kDa chaperonin (Protein Cpn60) (groEL protein).			

```

GN  GROL OR MOPA OR GROEL.
OS  Actinobacillus actinomycetemcomitans (Haemophilus
OC  actinomycetemcomitans).
OC  Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC  Actinobacillus.
OX  NCBI_TaxID=714;
RN  [1]
RP  SEQUENCE FROM N.A., AND SEQUENCE OF 1-39.
RC  STRAIN-Y4;
RX  MEDLINE=96017061; PubMed=7567064;
RA  Nakano T., Inai Y., Yamashita Y., Kusuzaki-Nagira T., Nagaoka S.,
RA  Okanashi N., Koga T., Nishihara T.;
RT  Molecular and immunological characterization of a 64-kDa protein of
RT  Actinobacillus actinomycetemcomitans."
RL  Oral Microbiol. Immunol. 10:151-159(1995).
CC  -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
CC  PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CC  CONDITIONS (BY SIMILARITY).
CC  -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
CC  7 SUBUNITS (BY SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
-----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
CC  EMBL; D28817; BAA05977.1; -
DR  HSSP; P06139; 1GRL.
DR  InterPro; IPR001844; Chaprin_Cpn60.
DR  InterPro; IPR002423; Cpn60/TCF-1.
DR  Pfam; PF00118; cpn60_TCF1.1
DR  PRINTS; PR00298; CHAPERONIN60.
DR  PRINTS; PR00304; TCOMPLEXTCP1.
DR  PROSITE; PS00296; CHAPERONINS_CPN60; 1.
KW  Chaperone; ATP-binding.
FT  INIT_MET 0
FT  INT_MET 0
SQ  SEQUENCE 546 AA; 57312 MW; 6249E0E46334F6F6 CRC64;

Query Match 61.8%; Score 34; DB 1; Length 546;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOOYTD 9
   1 111111
   478 NAGTEOYGD 486

RESULT 3
CH60_ACTPL STANDARD; PRT; 546 AA.
AC P94166;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 60 kDa chaperonin (Protein Cpn60) (grol protein).
GN GROL OR MOPA OR GROEL.
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Actinobacillus.
OX NCBI_TaxID=715;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S 4074 / Serotype 1;
RX MEDLINE=97189570; PubMed=9037757;
RA Vezina G., Strols M., Clairoux N., Boissinot M.;
RT Cloning and characterization of the grol locus from Actinobacillus
RT pleuropneumoniae."
RL FEMS Microbiol. Lett. 147:11-16(1997).
CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND

```

```

CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CC CONDITIONS (BY SIMILARITY).
CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
CC 7 SUBUNITS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; U55016; AAB51437.1; -
DR HSSP; P06139; 1GRL.
DR InterPro; IPR001844; Chaprin_Cpn60.
DR InterPro; IPR002423; Cpn60/TCF-1.
DR Pfam; PF00118; cpn60_TCF1.1
DR PRINTS; PR00298; CHAPERONIN60.
DR PRINTS; PR00304; TCOMPLEXTCP1.
DR PROSITE; PS00296; CHAPERONINS_CPN60; 1.
KW Chaperone; ATP-binding.
FT INIT_MET 0
FT INT_MET 0
SQ SEQUENCE 546 AA; 57513 MW; E84BB72C9BD3DB56 CRC64;

Query Match 61.8%; Score 34; DB 1; Length 546;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOOYTD 9
   1 111111
   479 NAGTEOYGD 487

RESULT 4
CH60_PASMU STANDARD; PRT; 547 AA.
AC Q59687;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 60 kDa chaperonin (Protein Cpn60) (grol protein).
GN GROL OR MOPA OR GROEL OR PM107.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-P1059;
RX MEDLINE=96105224; PubMed=8529887;
RA Love B.C., Hansen L.M., Hirsch D.C.;
RT Cloning and sequence of the grolS heat-shock operon of Pasteurella
RT multocida."
RL Gene 166:179-180(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whitlam T.S., Kapur V.;
RA "Complete genomic sequence of Pasteurella multocida pm70."
RT Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CC CONDITIONS (BY SIMILARITY).
CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
CC 7 SUBUNITS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```


CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 DR EMBL: U30165; AAA84916.1; -
 DR EMBL: AE006151; AAK03191.1; -
 DR HSSP: P06139; IJON.
 DR InterPro: IPR001844; Chaprinin_Cpn60.
 DR InterPro: IPR002423; Cpn60/TCP-1.
 DR Pfam: PF00118; cpn60_TCP1.1.
 DR PRINTS: PR00298; CHAPERONIN60.
 DR PRINTS: PR00304; TCOMPLEXTCP1.
 DR PROSITE: PS00296; CHAPERONINS.CPN60; 1.
 KW Chaperone; ATP-binding; Complete proteome.
 FT CONFLICT 424 424 S->N (IN REF. 1).
 FT CONFLICT 464 464 I->V (IN REF. 1).
 SQ SEQUENCE 547 AA; 57291 MW; EBB95B9F235B55E CRC64;

Query Match 61.8%; Score 34; DB 1; Length 547;
 Best Local Similarity 66.7%; Pred. No. 33;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTOOYTD 9
 Db 479 NAGTEOYGD 487

RESULT 5
 CH60_HAEIN STANDARD; PRT; 548 AA.
 ID CH60_HAEIN
 AC P43733;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 60 kDa chaperonin (Protein Cpn60) (groEL protein).
 GN GROEL OR MOXA OR GROEL OR HI0543.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus
 NCBI_TaxId=727;
 (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KM20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghagen N.S.M.,
 Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 Venter J.C.;
 RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
 Rd.";
 RT Science 269:496-512(1995).
 CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
 CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
 CC CONDITIONS (BY SIMILARITY).
 CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
 CC 7 SUBUNITS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 DR EMBL: U32736; AAC22201.1; -

DR HSSP: P06139; IJON.
 DR TRIGR: HI0543; -
 DR InterPro: IPR001844; Chaprinin_Cpn60.
 DR InterPro: IPR002423; Cpn60/TCP-1.
 DR Pfam: PF00118; cpn60_TCP1.1.
 DR PRINTS: PR00298; CHAPERONIN60.
 DR PRINTS: PR00304; TCOMPLEXTCP1.
 DR PROSITE: PS00296; CHAPERONINS.CPN60; 1.
 KW Chaperone; ATP-binding; Complete proteome.
 SQ SEQUENCE 548 AA; 57577 MW; CA4066AC1862159 CRC64;

Query Match 61.8%; Score 34; DB 1; Length 548;
 Best Local Similarity 66.7%; Pred. No. 33;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTOOYTD 9
 Db 480 NAGTEOYGD 488

RESULT 6
 C3BB_BACTU STANDARD; PRT; 652 AA.
 ID C3BB_BACTU
 AC Q06117; Q45717;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pesticidal crystal protein cry3Bb (insecticidal delta-endotoxin
 DE CryIIIB(b)) (Crystalline entomocidal protoxin) (74 kDa crystal
 DE protein).
 GN CRY3BB OR CRYIIIB(B) OR CRYIIIB2.
 OS Bacillus thuringiensis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 NCBI_TaxId=1428;
 (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=EG4961;
 RX MEDLINE=93119147; PubMed=1476436;
 RA Donovan W.P., Rupar M.J., Stanley A.C., Malvar T., Gawron-Burke M.C.,
 Johnson T.B.;
 RT "Characterization of two genes encoding Bacillus thuringiensis
 RT insecticidal crystal proteins toxic to Coleoptera species.";
 RL Appl. Environ. Microbiol. 58:3921-3927(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-18635 / EG5144;
 RA Donovan W.P., Rupar M.J., Stanley A.C.;
 RT "Bacillus thuringiensis cryIIIC, (b) protein toxic to coleopteran
 RT insects.";
 RL Patent number US5378625, 03-JAN-1995.
 CC -1- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
 CC EPITHELIAL CELLS OF COLEOPTERA. HAS MODERATE LEVEL OF TOXICITY TO
 CC SOUTHERN CORN ROOTWORM.
 CC -1- SUBUNIT: MONOMER.
 CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
 CC SPORELIATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART OF
 CC THE SPORE COAT.
 CC -1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE N-
 CC TERMINUS.
 CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 DR EMBL: M89794; AAA22334.1; -
 DR EMBL: U31633; AAY74198.1; -
 DR HSSP: P07130; IDLC.
 DR InterPro: IPR001178; Endotoxin.

DR Pfam: PF00555; endotoxin; 1.
 KW Toxin; Sporulation.
 FT VARIANT 21 21 Q -> P (IN STRAIN EG5144).
 FT VARIANT 97 97 N -> D (IN STRAIN EG5144).
 FT VARIANT 289 289 I -> V (IN STRAIN EG5144).
 FT VARIANT 352 352 F -> S (IN STRAIN EG5144).
 FT VARIANT 417 419 VYL -> IYF (IN STRAIN EG5144).
 FT VARIANT 451 451 S -> G (IN STRAIN EG5144).
 FT VARIANT 590 590 L -> I (IN STRAIN EG5144).
 FT VARIANT 600 600 K -> I (IN STRAIN EG5144).
 FT VARIANT 624 624 K -> T (IN STRAIN EG5144).
 SQ SEQUENCE 652 AA; 74385 MW; 63048332CDE8CC6 CRC64;
 Query Match 61.8%; Score 34; DB 1; Length 652;
 Best Local Similarity 87.5%; Pred. No. 39;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 KGTQOYTD 9
 DB 235 KLTQOYTD 242
 RESULT 7
 C3BA_BACTO STANDARD; PRT; 659 AA.
 ID C3BA_BACTO
 AC P17969;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 30-NOV-2000 (Rel. 39, Last annotation update)
 DE Pesticidal crystal protein cry3Ba (Insecticidal delta-endotoxin
 DE CryIIIBa) (Crystalline entomocidal protoxin) (75 kDa crystal
 DE protein).
 GN CRY3BA OR CRYIIIB(A) OR CRYIIIB.
 GN Bacillus thuringiensis (subsp. tolworthi).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1442;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=43F;
 RX MEDLINE=90206811; PubMed=2320431;
 RA Sick A., Gaertner F.H., Wong A.;
 RT Nucleotide sequence of a coleopteran-active toxin gene from a new
 RT isolate of Bacillus thuringiensis subsp. tolworthi.;
 RL Nucleic Acids Res. 18:1305-1305(1990).
 CC -1 FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
 CC EPIITHELIAL CELLS OF COLEOPTERA.
 CC -1 DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
 CC SPOULATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
 CC OF THE SPORE COAT.
 CC -1 MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE
 CC N-TERMINUS.
 CC -1 SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X17123; CAA34983.1; -
 CC EMBL: A07234; CAA00645.1; -
 CC PIR: S10228; S10228.
 CC HSSP: P07130; 1DLG.
 CC InterPro: IPR001178; Endotoxin.
 DR Pfam: PF00555; endotoxin; 1.
 KW Toxin; Sporulation.
 SQ SEQUENCE 659 AA; 75159 MW; 5A5B214FF8418CA CRC64;
 Query Match 61.8%; Score 34; DB 1; Length 659;
 Best Local Similarity 87.5%; Pred. No. 40;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYTD 9
 DB 243 KLTQOYTD 250
 RESULT 8
 CLPB_SYNY3 STANDARD; PRT; 872 AA.
 ID CLPB_SYNY3
 AC P74361;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE CLPB protein.
 DE CLPB OR SLR1641.
 GN Synechocystis sp. (strain PCC 6803).
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Nartuo K.,
 RA Okumura S., Shimizu S., Takeuchi C., Wada T., Watanabe A.,
 RA Yamada M., Yasuda M., Tabata S.;
 RT *Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 CC -1 SUBUNIT: THOUGHT TO BE AN ATPASE SUBUNIT OF AN INTRACELLULAR
 CC ATP-DEPENDENT PROTEASE (BY SIMILARITY).
 CC -1 SIMILARITY: BELONGS TO THE CLPB/CLPB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: D90914; BAA18456.1; -
 CC InterPro: IPR003593; AAA_ATPase.
 CC InterPro: IPR003959; AAA_ATPase_cent.
 CC InterPro: IPR001270; Chaprinin_clpA/B.
 CC InterPro: IPR004176; Clp_N.
 DR Pfam: PF00004; AAA; 2.
 DR Pfam: PF02861; Clp_N; 2.
 DR PRINTS: PR00300; CLP_PROTASEA.
 DR ProDom: PD000739; GSPIL_E; 1.
 DR SMART: SM00382; AAA; 2.
 DR PROSITE: PS00870; CLPB_1; 1.
 DR PROSITE: PS00871; CLPB_2; 1.
 KW Chapterone; ATP-binding; Repeat; Complete proteome.
 FT DOMAIN 163 411 I.
 FT DOMAIN 537 728 II.
 FT NP_BIND 208 215 ATP (POTENTIAL).
 FT NP_BIND 611 618 ATP (POTENTIAL).
 SQ SEQUENCE 872 AA; 98122 MW; 7EAA8486C7D8D4D5 CRC64;
 Query Match 61.8%; Score 34; DB 1; Length 872;
 Best Local Similarity 66.7%; Pred. No. 54;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 2 KGTQOYTDQ 10
 DB 146 KGTQOYTDQ 154
 RESULT 9
 GSHU_CAEEL STANDARD; PRT; 224 AA.
 ID GSHU_CAEEL

AC 095003;
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE Glutathione peroxidase precursor (EC 1.11.1.9).
 GN C11E4.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 NC NCB1_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Morimoto B.;
 RL Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: 2 glutathione + H(2)O(2) -> oxidized
 CC -1- SUBCELLULAR LOCATION: Extracellular (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE GLUTATHIONE PEROXIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; 281015; CAB02655.1; -
 DR HSSP; P00435; 1GPI.
 DR WormPep; C11E4.2; CE08102.
 DR InterPro; IPR000889; Glut_peroxidase.
 DR Pfam; PF00255; GSHpx; 1.
 DR PRINTS; PR01011; GLUTATHIONE_PEROXIDASE.
 DR PROSITE; PS00460; GLUTATHIONE_PEROXID_1; 1.
 DR PROSITE; PS00763; GLUTATHIONE_PEROXID_2; 1.
 DR Oxioreductase; Peroxidase; Signal.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 224 GLUTATHIONE PEROXIDASE.
 FT ACT_SITE 73 73 BY SIMILARITY.
 FT CARBOHYD 38 38 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 38 38
 SQ SEQUENCE 224 AA; 25556 MW; F02D055246DEZFI CRC64;
 Query Match 60.0%; Score 33; DB 1; Length 224;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Y 4 TQOYTD 9
 Db 76 TQOYTD 81
 RESULT 10
 ID TRY3_SALSA STANDARD; PRT; 238 AA.
 AC P35033;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trypsin III precursor (EC 3.4.21.4) (Fragment).
 OS Salmo salar (Atlantic salmon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
 NC NCB1_TaxID=8030;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=96035908; PubMed=7556223;
 RA Male R., Lorens J.B., Smals A.O., Torrisen K.R.;
 RT "Molecular Cloning and Characterization of anionic and cationic
 RT variants of trypsin from Atlantic salmon.";
 RL Eur. J. Biochem. 232:677-685(1995).

CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X70074; CAA49679.1; -
 DR PIR; S31779; S31779.
 DR HSSP; P00763; IDPO.
 DR MEROPS; S01.151; -
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS50240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 KW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
 KW Multigene family.
 FT NON_TER 1 1
 FT SIGNAL 1 7
 FT PROPEP 8 15 POTENTIAL.
 FT CHAIN 16 238 ACTIVATION PEPTIDE.
 FT ACT_SITE 55 55 TRYPSIN III.
 FT ACT_SITE 99 99 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 192 192 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 22 152 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 40 56 BY SIMILARITY.
 FT DISULFID 124 225 BY SIMILARITY.
 FT DISULFID 131 198 BY SIMILARITY.
 FT DISULFID 163 177 BY SIMILARITY.
 FT DISULFID 188 212 BY SIMILARITY.
 FT SITE 186 186 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 SQ SEQUENCE 238 AA; 25389 MW; AE799B80E8393023 CRC64;
 Query Match 60.0%; Score 33; DB 1; Length 238;
 Best Local Similarity 55.6%; Pred. No. 21;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NKGTQYTD 9
 Db 73 NKGTEQFD 81
 RESULT 11
 ID TRY2_XENLA STANDARD; PRT; 244 AA.
 AC P70059;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trypsin precursor (EC 3.4.21.4).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodidae; Xenopus.
 NC NCB1_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang K., Lytle L., Gan L., Hood L.E.;
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U72330; AAB1724.1; -
 DR HSP: F00763; IDPO.
 DR MEROPS: S01.258; -
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser-protease_Try.
 DR Pfam: PF00089; trypsin: 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; tryp-Spec: 1.
 DR PROSITE: PS00240; TRYPsin_DOM: 1.
 DR PROSITE: PS00134; TRYPsin_HIS: 1.
 DR PROSITE: PS00135; TRYPsin_SER: 1.
 KW Hydrolase: Serine protease; Digestion; Zymogen; Signal;
 KW Multigene family.
 FT SIGNAL 1 15 BY SIMILARITY.
 FT PROPEP 16 21 ACTIVATION PEPTIDE (BY SIMILARITY).
 FT CHAIN 22 244 TRYPsin.
 FT ACT_SITE 61 61 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 105 105 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 198 198 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 28 158 BY SIMILARITY.
 FT DISULFID 46 62 BY SIMILARITY.
 FT DISULFID 130 231 BY SIMILARITY.
 FT DISULFID 137 204 BY SIMILARITY.
 FT DISULFID 169 183 BY SIMILARITY.
 FT DISULFID 194 218 BY SIMILARITY.
 FT SITE 192 192 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 FT SEQUENCE 244 AA; 26079 MW; C63F29CB330B323 CRC64;
 Query Match 60.0%; Score 33; DB 1; Length 244;
 Best Local Similarity 55.6%; Pred. No. 22;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 NKGTQOYTD 9
 Db 79 NECTROFID 87
 RESULT 12
 ID UBIG_XYLFA STANDARD; PRT: 246 AA.
 AC O9PAM5;
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 3-demethylubiquinone-9-3-methyltransferase (EC 2.1.1.64) (3,4-
 dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHMB
 methyltransferase).
 GN UBIG OR XF2471.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xylella.
 ON NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9a5c;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Relnach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvares R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Bioness M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C., Ferro J.A.,
 RA Figueira J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohnsels J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuzama E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,

RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchino M.H.,
 RA Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Melandris J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*."
 RL Nature 406:151-159(2000).
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
 CC demethylubiquinone-9 - S-adenosyl-L-homocysteine + ubiquinone-9.
 CC -1- PATHWAY: ubiquinone biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE UBIG/COQ3 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE004055; AAF85269.1; -
 DR InterPro: IPR001601; Methyltransf.
 DR InterPro: IPR000501; SAM_bind.
 KW Ubiquinone biosynthesis; Transferase; Methyltransferase;
 KW Complete proteome.
 SQ SEQUENCE 246 AA; 27121 MW; 5212107D63633D3F CRC64;
 Query Match 60.0%; Score 33; DB 1; Length 246;
 Best Local Similarity 75.0%; Pred. No. 22;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Oy 2 KGTQOYTD 9
 Db 182 KGTQOYTD 189
 RESULT 13
 ID IFRH_TOBAC STANDARD; PRT: 310 AA.
 AC P52579;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Isoflavone reductase homolog A622 (EC 1.3.1.-).
 DE Nicotiana tabacum (Common tobacco).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 ON NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Burley 21; TISSUE=Root;
 RX MEDLINE=94312878; PubMed=8038607;
 RA Hibl N., Higashiguchi S., Hashimoto T., Yamada Y.;
 RT "Gene expression in tobacco low-nicotine mutants".
 RT Plant Cell 6:723-735(1994)
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: STRONG, TO ISOFLAVONE REDUCTASE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

```

CC      or send an email to license@lsb-sib.ch).
CC      -----
CC      EMBL; U66345; AAC49697.1; -.
CC      DR EMBL; AC006932; AAF22902.1; ALT_SEQ.
CC      DR InterPro: IPR001580; Calreticulin.
CC      DR InterPro: IPR000886; ER_target.
CC      DR Pfam; PF00262; calreticulin.1.
CC      DR PRINTS; PR00626; CALRETICULIN.
CC      DR ProDom; PD001866; Calreticulin.1.
CC      DR PROSITE; PS00014; ER_TARGET.1.
CC      DR PROSITE; PS00803; CALRETICULIN_1.1.
CC      DR PROSITE; PS00804; CALRETICULIN_2.1.
CC      DR PROSITE; PS00805; CALRETICULIN_REPEAT; FALSE NEG.
CC      KW Endoplasmic reticulum; Calcium-binding; Repeat; signal; Glycoprotein;
CC      MultiGene family.
CC      FT SIGNAL 1 28 POTENTIAL.
CC      FT CHAIN 29 424 CALRETICULIN_3.
CC      FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC      FT SITE 421 424 PREVENT SECRETION FROM ER (POTENTIAL).
CC      FT CONFLICT 279 279 F -> S (IN REF. 2).
CC      SQ SEQUENCE 424 AA; 49904 MW; 650E0AE8342F0B97 CRC64;
CC
CC      Query Match 60.0%; Score 33; DB 1; Length 424;
CC      Best Local Similarity 66.7%; Pred. No. 39;
CC      Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
CC
CC      Oy 1 NKG2OQYTD 9
CC          ||| |||
CC      Db 72 NKG1OTYND 80
CC
CC      RESULT 15
CC      6PGD_CITAM
CC      ID 6PGD_CITAM STANDARD; PRT; 445 AA.
CC      AC P41581.
CC      DT 01-NOV-1995 (Rel. 32, Created)
CC      DT 01-NOV-1995 (Rel. 32, Last sequence update)
CC      DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC      DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)
CC      DE (Fragment).
CC      GN GND.
CC      OS Citrobacter amalonaticus.
CC      OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC      Citrobacter.
CC      OX NCBI_TaxID=35703;
CC      [1]
CC      RP SEQUENCE FROM N.A.
CC      RC STRAIN=CT28;
CC      RX MEDLINE=95024018; PubMed=7937867;
CC      RA Nelson K., Selsander R.K.;
CC      RT "Intergenic transfer and recombination of the 6-phosphogluconate
CC      dehydrogenase gene (gnd) in enteric bacteria.",
CC      RL Proc. Natl. Acad. Sci. U.S.A. 91:10227-10231(1994).
CC      CC CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) -> D-ribulose
CC      5-phosphate + CO(2) + NADPH.
CC      -! PATHWAY: Hexose monophosphate shunt.
CC      -! SIMILARITY: BELONGS TO THE 6-PHOSPHOGLUCONATE DEHYDROGENASE
CC      FAMILY.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement. See http://www.isb-sib.ch/annouce/.
CC      or send an email to license@lsb-sib.ch).
CC
CC      EMBL; U14426; AAC43773.1; -.
CC      HSSP; P00349; 2PGD.
CC      DR InterPro: IPR001744; 6PGD.
CC      Pfam; PF00393; 6PGD.1.
CC      Pfam; PF03446; NAD_binding-2; 1.
CC      TRIGRAMS; TIGR00873; gnd; 1.
CC      DR.

```

DR PROSITE: PS00461; 6PGD; 1.
 KW Gluconate utilization; Oxidoreductase; Pentose shunt; NADP.
 FT NON_TER 1 1
 FT NON_TER 445 445
 SO SEQUENCE 445 AA; 48887 MW; A3771DRC9678EEF1 CRC64;
 Query Match 60.0%; Score 33; DB 1; Length 445;
 Best Local Similarity 60.0%; Pred. No. 41;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NKGTQGYTDQ 10
 DB 248 NKGTGKWTQ 257

Search completed: December 11, 2002, 15:29:02
 Job time : 12 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 15:27:52 : Search time 16 Seconds
(without alignments)
60.084 Million cell updates/sec

Title: US-09-991-433-5
Perfect score: 55
Sequence: 1 NKGTOQYTDQ 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	781	1	VCPY19 coat protein VP1 -
2	38	69.1	3971	2	mycosubtilin synth
3	37	67.3	338	2	CCAAr enhancer bin
4	35	63.6	169	2	prophage p13 prote
5	35	63.6	415	2	processing protein
6	34	61.8	106	2	conserved hypotet
7	34	61.8	375	2	gene 118 protein -
8	34	61.8	466	2	probable beta-keto
9	34	61.8	547	2	heat-shock protein
10	34	61.8	548	2	chaperonin groEL -
11	34	61.8	652	2	paraportal crystal
12	34	61.8	659	2	endopeptidase C1P
13	34	61.8	872	2	probable adhesin Z
14	34	61.8	1588	2	hypothetical prote
15	34	61.8	1588	2	hypothetical prote
16	33	60.0	163	2	hypothetical prote
17	33	60.0	223	2	hypothetical prote
18	33	60.0	224	2	hypothetical prote
19	33	60.0	238	2	hypothetical prote
20	33	60.0	246	2	tryptin (EC 3.4.21
21	33	60.0	310	2	3-demethylubiquino
22	33	60.0	375	2	2'-hydroxysiflavo
23	33	60.0	375	2	hypothetical prote
24	33	60.0	443	2	hypothetical prote
25	33	60.0	445	2	phosphogluconate d
26	33	60.0	445	2	phosphogluconate d
27	33	60.0	445	2	phosphogluconate d
28	33	60.0	445	2	phosphogluconate d
29	33	60.0	445	2	phosphogluconate d

30	33	60.0	445	2	140709 phosphogluconate d
31	33	60.0	468	1	DEECC phosphogluconate d
32	33	60.0	468	1	S04397 phosphogluconate d
33	33	60.0	468	2	162463 phosphogluconate d
34	33	60.0	468	2	162465 phosphogluconate d
35	33	60.0	468	2	D56146 phosphogluconate d
36	33	60.0	468	2	184555 phosphogluconate d
37	33	60.0	468	2	141249 phosphogluconate d
38	33	60.0	468	2	141250 phosphogluconate d
39	33	60.0	468	2	F90982 phosphogluconate d
40	33	60.0	468	2	D85828 gluconate-6-phosph
41	33	60.0	468	2	AE0765 phosphogluconate d
42	33	60.0	469	2	A10187 phosphogluconate d
43	33	60.0	474	2	E95043 hypothetical prote
44	33	60.0	474	2	G97913 phosphogluconate d
45	33	60.0	498	2	B90456 hypothetical prote

ALIGNMENTS

```

RESULT 1
VCPY19
coat protein VP1 - parvovirus B19 (strain Au)
C:Species: parvovirus B19
A:Note: host Homo sapiens (man)
C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999
C:Accession: A24299
R:Shade, R.O.; Blundell, M.C.; Cotmore, S.F.; Tattersall, P.; Astell, C.R.
J. Virol. 58, 921-936, 1986
A:Title: Nucleotide sequence and genome organization of human parvovirus B19 isolated
A:Reference number: A24299; MUID:86200451; PMID:3701931
A:Accession: A24299
A:Molecule type: DNA
A:Residues: 1-781 <SHA>
A:Cross-references: EMBL:M1178; NID:9333375; PIDN:AAA66867.1; PID:9333377
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein

Query Match 100.0%; Score 55; DB 1; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.0061;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQ 10
Db 622 NKGTOQYTDQ 631

RESULT 2
T44806
mycosubtilin synthetase chain mycA [Imported] - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 17-Nov-2000
C:Accession: T44806
R:Duitman, E.H.; Hamoen, L.W.; Rembold, M.; Venema, G.; Seltz, H.; Saenger, W.; Berni
Proc. Natl. Acad. Sci. U.S.A. 96, 13294-13299, 1999
A:Title: The mycosubtilin synthetase of Bacillus subtilis ATCC6633: A multifunctiona
A:Reference number: 222848; MUID:20027541; PMID:10557314
A:Accession: T44806
A:Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3971 <DUI>
A:Cross-references: EMBL:AF184956; NID:96449053; PIDN:AAF08795.1; PID:96449055
A:Experimental source: strain ATCC6633
C:Genetics:
A:Gene: mycA
C:Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acetate-CoA lig
C:Keywords: antibiotic biosynthesis; carrier protein
F:56-548/Domain: acetate-CoA ligase homology <ACLI>
F:581-649/Domain: acyl carrier protein homology <ACPI>
F:650-1090/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>
F:1290-1361/Domain: acyl carrier protein homology <ACPI>
F:2408-2476/Domain: acyl carrier protein homology <ACPI>

```

F:2986-3427/Domain: acetate-CoA ligase homology <ACL2>
F:3445-3513/Domain: acyl carrier protein homology <ACP4>

Query Match 69.1%; Score 38; DB 2; Length 3971;

Best Local Similarity 66.7%; Pred. No. 1;e+02;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KGTQOYTDQ 10

DB 1496 KGSQOYTD 1504

RESULT 3

CCNA7 enhancer-binding protein - California sea hare

N:Alternate names: C/EBP

C:Species: Aplysia californica (California sea hare)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000

C:Accession: A53066

A:Accession: A53066

A:Reference number: A53066; MUID:94185169; PMID:8137425

A:Title: C/EBP is an immediate-early gene required for the consolidation of long-term fa

A:Reference number: A53066; MUID:94185169; PMID:8137425

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-338 <ALB>

A:Cross-references: GB:U00994; NID:g392976; PIDN:AA18286.1; PID:g487963

Query Match 67.3%; Score 37; DB 2; Length 338;

Best Local Similarity 66.7%; Pred. No. 11;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYTDQ 10

DB 258 KGTQOYVDK 266

RESULT 4

prophage p13 protein 39 [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

C:Species: Lactococcus lactis subsp. lactis

C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C:Accession: D86801

R:Bolotin, A.; Mincer, P.; Manger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrlich

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss

A:Reference number: A86625; MUID:2125186; PMID:1137471

A:Accession: D86801

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-169 <STO>

A:Cross-references: GB:AE005176; PID:g12724400; PIDN:AAK0510.1; GSPDB:GND0146

A:Experimental source: strain IL1403

C:Genetics: p1339

Query Match 63.6%; Score 35; DB 2; Length 169;

Best Local Similarity 60.0%; Pred. No. 13;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10

DB 60 NNSQOYADQ 69

RESULT 5

processing proteinase homolog ymfH - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000

C:Accession: G69885

R:Kuncl, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Allioni, G.; Azevedo, V.; Berte

C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallazzi, A.; Gal

lechi, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau

Y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portele

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Toononi, A.; Tosato, V.; Uchiya

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: G69885

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-415 <KUN>

A:Cross-references: GB:299112; GB:AL009126; NID:g2633902; PIDN:CA13559.1; PID:g26340

A:Experimental source: strain 168

C:Genetics:

A:Gene: ymfH

C:Superfamily: mitochondrial processing peptidase alpha chain

Query Match 63.6%; Score 35; DB 2; Length 415;

Best Local Similarity 60.0%; Pred. No. 35;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10

DB 208 NQCKRPTDQ 217

RESULT 6

conserved hypothetical protein SA0887 [imported] - Staphylococcus aureus (strain N315

C:Species: Staphylococcus aureus

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C:Accession: A89872

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O

ma, A.; Mizutani, U.; Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K

C: Sliba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: A89872

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-106 <KUR>

A:Cross-references: GB:BA000018; PID:g13700836; PIDN:BA842132.1; GSPDB:GND0149

A:Experimental source: strain N315

C:Genetics:

A:Gene: SA0887

Query Match 61.8%; Score 34; DB 2; Length 106;

Best Local Similarity 75.0%; Pred. No. 12;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYTD 9

DB 45 KGTQOYVN 52

RESULT 7

gene 118 protein - Lactococcus phage b11170

C:Species: Lactococcus phage b11170

C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 04-Mar-2000

C:Accession: T03325

R:Curry-Coy, A.M.; Cesselin, B.; Commissaire, J.; Anba, J.; Kyriakidis, S.; Chopin

submitted to the EMBL Data Library, June 1997

A:Description: Sequence and organization of the lactococcal isometric b11170 phage ge

A:Reference number: Z14903
A:Accession: T03325
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-375 <CRU>
A:Cross-references: EMBL:AF009630; NID:g3282260; PIDN:AAC27197.1; PID:g3282278
C:Genetics:
A:Gene: 118
C:Superfamily: Lactococcus phage DIL170 gene 118 protein

Query Match 61.8%; Score 34; DB 2; Length 375;
Best Local Similarity 66.7%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 NKGTQOYTD 9
DB 250 NKATDYTD 258

RESULT 8

probable beta-ketoacyl-CoA synthase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: D84906
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanKen, S.E.; Unayam, L.; Tallon, L.; eus, D.; Nienhan, W.C.; White, O.; Eissen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84906
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-466 <STO>
A:Cross-references: GB:AE002093; NID:g3831446; PIDN:AAC69929.1; GSPDB:GN00139
C:Genetics:
A:Gene: AT2946720
A:Map position: 2

Query Match 61.8%; Score 34; DB 2; Length 466;
Best Local Similarity 66.7%; Pred. No. 65;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 KGTQOYTDQ 10
DB 86 KGAOLYTD 94

RESULT 9

heat-shock protein GroEL - Pasteurella multocida
C:Species: Pasteurella multocida
C:Date: 08-Feb-1996 #sequence_revision 19-Apr-1996 #text_change 26-Aug-1999
C:Accession: J04519
R:Love, B.C.; Hansen, L.M.; Hirsh, D.C.
Gene 166, 179-180, 1995
A:Title: Cloning and sequence of the groESL heat-shock operon of Pasteurella multocida.
A:Reference number: J04518; MUID:96105224; PMID:8529887
A:Accession: J04519
A:Molecule type: DNA
A:Residues: 1-547 <LOV>
A:Cross-references: GB:U30165; NID:g1144300; PIDN:AA84916.1; PID:g1144302
A:Experimental source: serotype A:3
C:Genetics:
A:Gene: groEL
C:Superfamily: chaperonin groEL
C:Keywords: heat shock; stress-induced protein

Query Match 61.8%; Score 34; DB 2; Length 547;
Best Local Similarity 66.7%; Pred. No. 77;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 NKGTQOYTD 9
DB 479 NAGTEQYGD 487

RESULT 10

chaperonin groEL - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 26-Aug-1999
C:Accession: C64076
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kiehlavag
R:Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: C64076
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-548 <TIGR>
A:Cross-references: GB:U32736; GB:L42023; NID:g1573519; PIDN:AAC22201.1; PID:g1573528
C:Genetics:
A:Gene: groEL
C:Superfamily: chaperonin groEL
C:Keywords: heat shock; molecular chaperone; stress-induced protein

Query Match 61.8%; Score 34; DB 2; Length 548;
Best Local Similarity 66.7%; Pred. No. 77;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 NKGTQOYTD 9
DB 480 NAGTEQYGD 488

RESULT 11

parasporal crystal protein cry3Bb1 - Bacillus thuringiensis
N:Alternate names: parasporal crystal protein cryIIIB2
C:Species: Bacillus thuringiensis
C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 01-Dec-2000
C:Accession: J39811
R:Donovan, W.P.; Rupar, M.J.; Slaney, A.C.; Malvar, T.; Gawron-Burks, M.C.; Johnson, A.
Appl. Environ. Microbiol. 58, 3921-3927, 1992
A:Title: Characterization of two genes encoding Bacillus thuringiensis insecticidal
A:Reference number: J39811; MUID:93119147; PMID:1476436
A:Accession: J39811
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-652 <RES>
A:Cross-references: GB:M89794; NID:g142729; PIDN:AA22334.1; PID:g142730
C:Genetics:
A:Gene: cryIIIB2
C:Superfamily: parasporal crystal protein

Query Match 61.8%; Score 34; DB 2; Length 652;
Best Local Similarity 87.5%; Pred. No. 94;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 KGTQOYTD 9
DB 235 KLTQOYTD 242

RESULT 12

parasporal crystal protein cry3Ba1 - Bacillus thuringiensis (fragment)
N:Alternate names: coleopter-an-active parasporal crystal protein; delta-endotoxin
C:Species: Bacillus thuringiensis
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 01-Dec-2000
C:Accession: S10228

R:Sick, A.; Gaertner, F.; Wong, A.
Nucleic Acids Res. 18, 1305, 1990
A:Title: Nucleotide sequence of a coleopteran-active toxin gene from a new isolate of Ba
A:Reference number: S10228; PMID:90206811; PMID:2320431
A:Accession: S10228
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-659 <SIC>
A:Cross-references: EMBL:X17123; NID:g40258; PIDN:CAA34983.1; PID:g40259
C:Genetics:
A:Gene: cryIIIB
C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin; toxin

Query Match 61.8%; Score 34; DB 2; Length 659;
Best Local Similarity 87.5%; Pred. No. 95;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 KGTQYTDQ 9
1:|||||
243 KLTQYTDQ 250

RESULT 13

S76197
endopeptidase Clp ATP-binding chain B1 - *Synechocystis* sp. (strain PCC 6803)
N:Alternate names: ATP-dependent Clp proteinase regulatory chain; protein slr1641
N:Contains: adenosinetriphosphatase (EC 3.6.1.3)
G:Species: *Synechocystis* sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 19-Jan-2001

C:Accession: S76197
R:Kaneho, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimp, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

9.
A:Reference number: S74322; PMID:97061201; PMID:8905231
A:Accession: S76197
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-872 <KAN>

A:Cross-references: EMBL:D90914; GB:AB001339; NID:g1653477; PIDN:BA18456.1; PID:g165354
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: clpB1
C:Function:

A:Description: allows clpP to hydrolyze polypeptides and proteins, probably by a chaperon
activity; ATP hydrolysis is required for Clp hydrolysis of proteins but not of smaller
superfamily: endopeptidase Clp ATP-binding chain
C:Keywords: ATP; hydrolyase; molecular chaperone; nucleotide binding; P-loop
F:208-215/Region: nucleotide-binding motif A (P-loop)
F:276-280/Region: nucleotide-binding motif B
F:511-618/Region: nucleotide-binding motif A (P-loop)
F:579-683/Region: nucleotide-binding motif B
F:214/Binding site: ATP (lys) #status predicted
F:617/Binding site: ATP (lys) #status predicted

Query Match 61.8%; Score 34; DB 2; Length 872;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 KGTQYTDQ 10
1:|||||
Db 146 RGTQYTDQ 154

RESULT 14

A86036
probable adhesin z5029 [imported] - *Escherichia coli* (strain O157:H7, substrain EDL933)
C:Species: *Escherichia coli*
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: A86036

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A:Reference number: A85480; PMID:21074935; PMID:11206551
A:Accession: A86036
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1588 <STO>
A:Cross-references: GB:AE005174; NID:g12518349; PIDN:AAG58749.1; GSPDB:GN00145; UWGP:

A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: z5029

Query Match 61.8%; Score 34; DB 2; Length 1588;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 TQOYTDQ 10
1:|||||
Db 1492 TKOYTDQ 1498

RESULT 15

H91188
probable adhesin ECS4480 [similarity] - *Escherichia coli* (strain O157:H7, substrain R
C:Species: *Escherichia coli*
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 02-Nov-2001

C:Accession: H91188
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and g
A:Reference number: A99629; PMID:21156231; PMID:11258796
A:Accession: H91188

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1588 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA837903.1; PID:g1363955; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECS4480

Query Match 61.8%; Score 34; DB 2; Length 1588;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 TQOYTDQ 10
1:|||||
Db 1492 TKOYTDQ 1498

Search completed: December 11, 2002, 15:30:00
Job time : 17 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 15:28:48 : Search time 11 Seconds
(without alignments)
14.766 Million cell updates/sec

Title: US-09-991-433-5
Perfect score: 55
Sequence: 1 NKG2GQYTDQ 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Jcal number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	61.8	872	9	US-10-047-260-38
2	33	60.0	175	2	US-10-062-254-234
3	33	60.0	467	10	US-09-875-573-17
4	33	60.0	467	10	US-09-875-573-19
5	33	60.0	467	10	US-09-875-573-21
6	33	60.0	468	10	US-09-815-242-10201
7	33	60.0	468	10	US-09-815-242-13822
8	33	60.0	468	10	US-09-875-573-23
9	33	60.0	468	10	US-09-875-573-25
10	33	60.0	468	10	US-09-875-573-27
11	33	60.0	468	10	US-09-875-573-29
12	33	60.0	468	10	US-09-875-573-31
13	33	60.0	468	10	US-09-875-573-33
14	33	60.0	468	10	US-09-875-573-35
15	33	60.0	468	10	US-09-875-573-37
16	33	60.0	468	10	US-09-875-573-39
17	33	60.0	468	10	US-09-875-573-41
18	33	60.0	468	10	US-09-875-573-43
19	33	60.0	473	10	US-09-815-242-5008

20	33	60.0	473	10	US-09-815-242-10478	Sequence 10478, A
21	33	60.0	481	10	US-09-815-242-13425	Sequence 13425, A
22	33	60.0	557	12	US-10-052-586-326	Sequence 326, App
23	31	56.4	239	10	US-09-910-071-15	Sequence 15, Appl
24	31	56.4	330	9	US-09-870-759-82	Sequence 82, Appl
25	31	56.4	428	10	US-09-347-331-10	Sequence 10, Appl
26	31	56.4	456	10	US-09-815-242-10400	Sequence 10400, A
27	31	56.4	456	10	US-09-815-242-11073	Sequence 11073, A
28	31	56.4	456	10	US-09-815-242-14064	Sequence 14064, A
29	31	56.4	458	10	US-09-815-242-11759	Sequence 11759, A
30	31	56.4	644	10	US-09-943-692-2	Sequence 2, Appl1
31	31	56.4	666	10	US-09-765-272-2	Sequence 2, Appl1
32	31	56.4	719	10	US-09-815-242-13423	Sequence 13423, A
33	31	56.4	778	10	US-09-815-242-10728	Sequence 10728, A
34	31	56.4	2012	9	US-09-808-602-68	Sequence 68, Appl1
35	30	54.5	98	10	US-09-811-384-2	Sequence 2, Appl1
36	30	54.5	116	10	US-09-910-059-22	Sequence 22, Appl1
37	30	54.5	131	10	US-09-864-761-41079	Sequence 41079, A
38	30	54.5	255	10	US-09-910-059-19	Sequence 19, Appl1
39	30	54.5	255	10	US-09-910-059-57	Sequence 57, Appl1
40	30	54.5	283	10	US-09-815-242-13476	Sequence 13476, A
41	30	54.5	310	9	US-09-989-350-10	Sequence 10, Appl1
42	30	54.5	310	9	US-09-840-707A-20	Sequence 20, Appl1
43	30	54.5	326	9	US-10-047-542-22	Sequence 22, Appl1
44	30	54.5	417	10	US-09-828-313-31	Sequence 31, Appl1
45	30	54.5	434	10	US-09-430-221-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-10-047-260-38
; Sequence 38, Application US/10047260
; Patent No. US20020164706A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Lisa
; APPLICANT: McCluskey, Michael
; APPLICANT: Larossa, Robert
; TITLE OR INVENTION: HIGH LEVEL PROMOTERS FROM CYANOBACTERIA
; FILE REFERENCE: CL1715 US NA
; CURRENT APPLICATION NUMBER: US/10/047,260
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/264,925
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 38
; LENGTH: 872
; TYPE: PRT
; ORGANISM: *Synechocystis* sp. strain PCC6803
US-10-047-260-38

Query Match 61.8%; Score 34; DB 9; Length 872;
Best Local Similarity 66.7%; Pred. No. 52;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 KGTQGYTDQ 10
Db 146 RGTQGYTDQ 154

RESULT 2
US-10-062-254-234
; Sequence 234, Application US/10062254
; Patent No. US20020138882A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Edgar B
; APPLICANT: Cahoon, Rebecca E
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Fany, Yiven
; APPLICANT: Hantke, Sabine S.
; APPLICANT: Lee, Jian-Ming

```

; APPLICANT: Li, Zhongsen
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Morgante, Michele
; APPLICANT: Niu, Xiping
; APPLICANT: Odell, Joan
; APPLICANT: Rafalski, Antoni
; APPLICANT: Sakai, Hajime
; APPLICANT: Zheng, Peizhong
; APPLICANT: Zhu, Qun
; TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/062,254
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 09/630,346
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/146511
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/156006
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/156899
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 60/157287
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 60/169767
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/171054
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/172958
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/171515
; PRIOR FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: 60/173535
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 375
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 234
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Trillium aestivum
; US-10-062-254-234

Query Match          60.0%; Score 33; DB 12; Length 175;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOY 8
    ||||:|
Db 77 NKGTRLXT 84

RESULT 3
US-09-875-573-17
; Sequence 17, Application US/09875573
; Patent No. US20020150902A1
; GENERAL INFORMATION:
; APPLICANT: Tarr, Phillip I.
; TITLE OF INVENTION: POLYMORPHIC LOCI THAT DIFFERENTIATE
; FILE REFERENCE: CHMED.001C1
; CURRENT APPLICATION NUMBER: US/09/875,573
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US99/29149
; PRIOR FILING DATE: 1999-12-08
; PRIOR FILING DATE: 1998-12-08
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-09-875-573-17
```

```

Query Match          60.0%; Score 33; DB 10; Length 467;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
    ||||:|
Db 258 NKGTGKWT SQ 267

RESULT 4
US-09-875-573-19
; Sequence 19, Application US/09875573
; Patent No. US20020150902A1
; GENERAL INFORMATION:
; APPLICANT: Tarr, Phillip I.
; TITLE OF INVENTION: POLYMORPHIC LOCI THAT DIFFERENTIATE
; FILE REFERENCE: CHMED.001C1
; CURRENT APPLICATION NUMBER: US/09/875,573
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US99/29149
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/111,493
; PRIOR FILING DATE: 1998-12-08
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-09-875-573-19

Query Match          60.0%; Score 33; DB 10; Length 467;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
    ||||:|
Db 258 NKGTGKWT SQ 267

RESULT 5
US-09-875-573-21
; Sequence 21, Application US/09875573
; Patent No. US20020150902A1
; GENERAL INFORMATION:
; APPLICANT: Tarr, Phillip I.
; TITLE OF INVENTION: POLYMORPHIC LOCI THAT DIFFERENTIATE
; FILE REFERENCE: CHMED.001C1
; CURRENT APPLICATION NUMBER: US/09/875,573
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US99/29149
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/111,493
; PRIOR FILING DATE: 1998-12-08
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-09-875-573-21

Query Match          60.0%; Score 33; DB 10; Length 467;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
    ||||:|
Db 258 NKGTGKWT SQ 267
```

```
RESULT 6
US-09-815-242-10201
; Sequence 10201, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlson, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10201
; LENGTH: 468
; TYPE: PRF
; ORGANISM: Escherichia coli
US-09-815-242-10201

Query Match          60.0%; Score 33; DB 10; Length 468;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOXYTDQ 10
||||:..|
DB 259 NKGTGKWT SQ 268

RESULT 7
US-09-815-242-13822
; Sequence 13822, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlson, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10201
; LENGTH: 468
; TYPE: PRF
; ORGANISM: Escherichia coli
US-09-815-242-10201
```

```
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 13822
;; LENGTH: 468
;; TYPE: PRF
;; ORGANISM: Salmonella typhi
US-09-815-242-13822

Query Match          60.0%; Score 33; DB 10; Length 468;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOXYTDQ 10
||||:..|
DB 259 NKGTGKWT SQ 268

RESULT 8
US-09-875-573-23
; Sequence 23, Application US/09875573
; Patent No. US20020150902A1
; GENERAL INFORMATION:
; APPLICANT: Tarr, Phillip I.
; TITLE OF INVENTION: POLYMORPHIC LOCI THAT DIFFERENTIATE
; FILE REFERENCE: CHMED.001C1
; CURRENT APPLICATION NUMBER: US/09/875,573
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US99/29149
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/111,493
; PRIOR FILING DATE: 1998-12-08
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 468
; TYPE: PRF
; ORGANISM: Escherichia coli
US-09-875-573-23

Query Match          60.0%; Score 33; DB 10; Length 468;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOXYTDQ 10
||||:..|
DB 259 NKGTGKWT SQ 268

RESULT 9
US-09-875-573-25
; Sequence 25, Application US/09875573
; Patent No. US20020150902A1
; GENERAL INFORMATION:
; APPLICANT: Tarr, Phillip I.
; TITLE OF INVENTION: POLYMORPHIC LOCI THAT DIFFERENTIATE
; FILE REFERENCE: CHMED.001C1
; CURRENT APPLICATION NUMBER: US/09/875,573
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US99/29149
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/111,493
; PRIOR FILING DATE: 1998-12-08
; NUMBER OF SEQ ID NOS: 43
```

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 25
LENGTH: 468
TYPE: PRT
ORGANISM: Escherichia coli
US-09-875-573-25

Query Match 60.0%; Score 33; DB 10; Length 468;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
||||:|
DB 259 NKGTGKWT SQ 268

RESULT 10
US-09-875-573-27
Sequence 27, Application US/09875573
Patent No. US20020150902A1
GENERAL INFORMATION:
APPLICANT: Tarr, Phillip I.

TITLE OF INVENTION: POLYMORPHIC LOCI THAT DIFFERENTIATE
FILE REFERENCE: CHMED.001C1
CURRENT APPLICATION NUMBER: US/09/875,573
CURRENT FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: PCT/US99/29149
PRIOR FILING DATE: 1999-12-08
PRIOR APPLICATION NUMBER: US 60/111,493
PRIOR FILING DATE: 1998-12-08
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 27
LENGTH: 468
TYPE: PRT
ORGANISM: Escherichia coli
US-09-875-573-27

Query Match 60.0%; Score 33; DB 10; Length 468;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
||||:|
DB 259 NKGTGKWT SQ 268

ULT 11
09-875-573-29
Sequence 29, Application US/09875573
Patent No. US20020150902A1
GENERAL INFORMATION:
APPLICANT: Tarr, Phillip I.
TITLE OF INVENTION: POLYMORPHIC LOCI THAT DIFFERENTIATE
FILE REFERENCE: CHMED.001C1
CURRENT APPLICATION NUMBER: US/09/875,573
CURRENT FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: PCT/US99/29149
PRIOR FILING DATE: 1999-12-08
PRIOR APPLICATION NUMBER: US 60/111,493
PRIOR FILING DATE: 1998-12-08
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 29
LENGTH: 468
TYPE: PRT
ORGANISM: Escherichia coli
US-09-875-573-29

Query Match 60.0%; Score 33; DB 10; Length 468;
Best Local Similarity 60.0%; Pred. No. 43;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 NKGTQOYTDQ 10
||||:|
DB 259 NKGTGKWT SQ 268

RESULT 12
US-09-875-573-31
Sequence 31, Application US/09875573
Patent No. US20020150902A1
GENERAL INFORMATION:
APPLICANT: Tarr, Phillip I.

TITLE OF INVENTION: POLYMORPHIC LOCI THAT DIFFERENTIATE
FILE REFERENCE: CHMED.001C1
CURRENT APPLICATION NUMBER: US/09/875,573
CURRENT FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: PCT/US99/29149
PRIOR FILING DATE: 1999-12-08
PRIOR APPLICATION NUMBER: US 60/111,493
PRIOR FILING DATE: 1998-12-08
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 31
LENGTH: 468
TYPE: PRT
ORGANISM: Escherichia coli
US-09-875-573-31

Query Match 60.0%; Score 33; DB 10; Length 468;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
||||:|
DB 259 NKGTGKWT SQ 268

RESULT 13
US-09-875-573-33
Sequence 33, Application US/09875573
Patent No. US20020150902A1
GENERAL INFORMATION:
APPLICANT: Tarr, Phillip I.
TITLE OF INVENTION: POLYMORPHIC LOCI THAT DIFFERENTIATE
FILE REFERENCE: CHMED.001C1
CURRENT APPLICATION NUMBER: US/09/875,573
CURRENT FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: PCT/US99/29149
PRIOR FILING DATE: 1999-12-08
PRIOR APPLICATION NUMBER: US 60/111,493
PRIOR FILING DATE: 1998-12-08
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 33
LENGTH: 468
TYPE: PRT
ORGANISM: Escherichia coli
US-09-875-573-33

Query Match 60.0%; Score 33; DB 10; Length 468;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
||||:|
DB 259 NKGTGKWT SQ 268

RESULT 14
US-09-875-573-35

```

; Sequence 35, Application US/09875573
; Patent No. US20020150902A1
; GENERAL INFORMATION:
; APPLICANT: Tarr, Phillip I.
; TITLE OF INVENTION: POLYMORPHIC LOCI THAT DIFFERENTIATE
; TITLE OF INVENTION: ESCHERICHIA COLI O157:H7 FROM OTHER STRAINS
; FILE REFERENCE: CHMED.001C1
; CURRENT APPLICATION NUMBER: US/09/875,573
; CURRENT FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US99/29149
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/111,493
; PRIOR FILING DATE: 1998-12-08
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Escherichia coli
; S-09-875-573-35

```

```

Query Match          60.0%; Score 33; DB 10; Length 468;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

OY      1  NKGTQOYTDQ 10
      |||| :|||
Db      259  NKGTGKWT SQ 268

```

```

RESULT 15
US-09-875-573-37
; Sequence 37, Application US/09875573
; Patent No. US20020150902A1
; GENERAL INFORMATION:
; APPLICANT: Tarr, Phillip I.
; TITLE OF INVENTION: POLYMORPHIC LOCI THAT DIFFERENTIATE
; TITLE OF INVENTION: ESCHERICHIA COLI O157:H7 FROM OTHER STRAINS
; FILE REFERENCE: CHMED.001C1
; CURRENT APPLICATION NUMBER: US/09/875,573
; CURRENT FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US99/29149
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/111,493
; PRIOR FILING DATE: 1998-12-08
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-09-875-573-37

```

```

Query Match          60.0%; Score 33; DB 10; Length 468;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

OY      1  NKGTQOYTDQ 10
      |||| :|||
Db      259  NKGTGKWT SQ 268

```

Search completed: December 11, 2002, 15:30:38
Job time : 12 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 15:28:08 ; Search time 15 Seconds
(without alignments)
19.615 Million cell updates/sec

Title: US-09-991-433-5
Perfect score: 55
Sequence: 1 NKGCTQYTDQ 10

Scoring table: BLOSUM62
Gap 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents: AA*
1: /cgn2.6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2.6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2.6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2.6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2.6/ptodata/1/1aa/PCRTUS.COMB.pep:*
6: /cgn2.6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	395	4	US-08-856-841-13 Sequence 13, Appl
2	55	100.0	398	4	US-08-856-841-21 Sequence 21, Appl
3	55	100.0	415	4	US-08-856-841-20 Sequence 20, Appl
4	55	100.0	543	4	US-08-856-841-22 Sequence 22, Appl
5	35	63.6	415	4	US-09-554-999-2 Sequence 203, Appl
6	34	61.8	187	4	US-09-095-855-203 Sequence 203, Appl
7	34	61.8	187	4	US-09-205-426-203 Sequence 14, Appl
8	34	61.8	373	4	US-09-039-198A-14 Sequence 15, Appl
9	34	61.8	373	4	US-09-039-198A-15 Sequence 15, Appl
10	34	61.8	373	4	US-08-877-599-15 Sequence 15, Appl
11	34	61.8	373	4	US-09-267-574-14 Sequence 14, Appl
12	34	61.8	373	4	US-09-267-574-15 Sequence 13, Appl
13	34	61.8	387	2	US-08-486-839-6 Sequence 6, Appl
14	34	61.8	387	2	US-09-151-011-6 Sequence 6, Appl
15	34	61.8	387	2	US-09-343-623-6 Sequence 4, Appl
16	34	61.8	466	2	US-08-486-839-4 Sequence 4, Appl
17	34	61.8	466	3	US-08-151-011-4 Sequence 2, Appl
18	34	61.8	466	3	US-09-039-198A-2 Sequence 2, Appl
19	34	61.8	466	4	US-09-343-623-4 Sequence 4, Appl
20	34	61.8	466	4	US-08-877-599-2 Sequence 2, Appl
21	34	61.8	466	4	US-09-267-574-2 Sequence 2, Appl
22	34	61.8	493	3	US-08-996-441B-70 Sequence 70, Appl
23	34	61.8	493	3	US-08-993-722A-70 Sequence 70, Appl
24	34	61.8	493	3	US-08-993-170A-70 Sequence 70, Appl
25	34	61.8	493	3	US-08-993-775B-70 Sequence 70, Appl
26	34	61.8	651	1	US-08-315-468-6 Sequence 6, Appl
27	34	61.8	651	1	US-08-315-468-6 Sequence 6, Appl

28	34	61.8	651	3	US-08-996-441B-52 Sequence 52, Appl
29	34	61.8	651	3	US-08-996-441B-56 Sequence 56, Appl
30	34	61.8	651	3	US-08-996-441B-58 Sequence 58, Appl
31	34	61.8	651	3	US-08-993-722A-52 Sequence 52, Appl
32	34	61.8	651	3	US-08-993-722A-56 Sequence 56, Appl
33	34	61.8	651	3	US-08-993-722A-58 Sequence 58, Appl
34	34	61.8	651	3	US-08-993-170A-52 Sequence 52, Appl
35	34	61.8	651	3	US-08-993-170A-56 Sequence 56, Appl
36	34	61.8	651	3	US-08-993-170A-58 Sequence 58, Appl
37	34	61.8	651	3	US-08-993-775B-52 Sequence 52, Appl
38	34	61.8	651	3	US-08-993-775B-56 Sequence 56, Appl
39	34	61.8	651	3	US-08-993-775B-58 Sequence 58, Appl
40	34	61.8	651	4	US-07-941-650A-4 Sequence 4, Appl
41	34	61.8	652	3	US-08-996-441B-2 Sequence 2, Appl
42	34	61.8	652	3	US-08-996-441B-4 Sequence 4, Appl
43	34	61.8	652	3	US-08-996-441B-6 Sequence 6, Appl
44	34	61.8	652	3	US-08-996-441B-8 Sequence 8, Appl
45	34	61.8	652	3	US-08-996-441B-12 Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-856-841-13
Sequence 13, Application US/08856841

Patent No. 6274307

GENERAL INFORMATION:

APPLICANT: ERWIN SOUTSCHEK

APPLICANT: MANFRED MOTZ

TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES

TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY

STREET: 99 PARK AVENUE

CITY: NEW YORK

STATE: NY

COUNTRY: USA

ZIP: 10016

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" FLOPPY DISC

COMPUTER: AT&T - IBM COMPATIBLE

OPERATING SYSTEM: MS-DOS Version 6.2

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/856,841

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/214,658

FILING DATE: 16-MARCH-1994

APPLICATION NUMBER: US 07/917,096

FILING DATE: 4-AUGUST-1992

APPLICATION NUMBER: PCT/DE91/00106

FILING DATE: 8-FEBRUARY-1991

APPLICATION NUMBER: DE40038262

FILING DATE: 8-FEBRUARY-1990

ATTORNEY/AGENT INFORMATION:

NAME: ROBINSON, WILLIAM R.

REGISTRATION NUMBER: 27,224

REFERENCE/DOCKET NUMBER: LKR-9222-A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 697-3355

TELEFAX: (212) 557-5635

TELEX: NONE

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 395

TYPE: AMINO ACID

TOPOLOGY: LINEAR

MOLECULE TYPE:

DESCRIPTION: PEPTIDE

1 HYPOTHETICAL: N/A
2 ANTI-SENSE: N/A
3 FRAGMENT TYPE: INTERNAL
4 ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
5 ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
6 IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
7 POSITION IN GENOME: N/A
8 FEATURE:
9 NAME/KEY: N/A
10 LOCATION: N/A
11 IDENTIFICATION METHOD: amino acid analysis and
12 OTHER INFORMATION: mass spectrometry
13 PUBLICATION INFORMATION:
14 AUTHORS: COSSART, Y.E.
15 AUTHORS: FIELD, A.M.
16 AUTHORS: CANT, B.
17 AUTHORS: WIDDOWS, D.
18 TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
19 JOURNAL: LANCET
20 VOLUME: 1
21 ISSUE:
22 PAGES: 72 - 73
23 DATE: 1975
24 DOCUMENT NUMBER:
25 FILING DATE:
26 PUBLICATION DATE:
27 RELEVANT RESIDUES IN SEQ ID NO: 13:
28 PUBLICATION INFORMATION:
29 AUTHORS: MANIATIS, T.
30 AUTHORS: FRITSCH, E.F.
31 AUTHORS: SAMBROOK, J.
32 TITLE: MOLECULAR CLONING
33 JOURNAL: COLD SPRING HARBOR, NY
34 VOLUME:
35 ISSUE:
36 PAGES:
37 DATE: 1982
38 DOCUMENT NUMBER:
39 FILING DATE:
40 PUBLICATION DATE:
41 RELEVANT RESIDUES IN SEQ ID NO: 13:
42 PUBLICATION INFORMATION:
43 AUTHORS: SMITH, D.B.
44 AUTHORS: JOHNSON, K.S.
45 TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES
46 TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH
47 JOURNAL: GLUTATHIONE S. TRANSFERASE
48 VOLUME:
49 ISSUE: 67
50 PAGES: 31 - 40
51 DATE: 1988
52 DOCUMENT NUMBER:
53 FILING DATE:
54 PUBLICATION DATE:
55 RELEVANT RESIDUES IN SEQ ID NO: 13:
56 US-08-856-841-13
57
58 Query Match 100.0%; Score 55; DB 4; Length 395;
59 Best Local Similarity 100.0%; Fred. No. 0.0037;
60 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
61
62 QY 1 NKGTQOYTDO 10
63 1111111111
64 DB 256 NKGTQOYTDO 265
65
66 RESULT 2
67 US-08-856-841-21
68 Sequence 21, Application US/08856841
69 Patent No. 6274307
70 GENERAL INFORMATION:

1 APPLICANT: ERWIN SOUTSCHER
2 APPLICANT: MANFRED MOTZ
3 TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
4 TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
5 NUMBER OF SEQUENCES: 28
6 CORRESPONDENCE ADDRESS:
7 ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
8 STREET: 99 PARK AVENUE
9 CITY: NEW YORK
10 STATE: NY
11 COUNTRY: USA
12 ZIP: 10016
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE: 3.5" FLOPPY DISC
15 COMPUTER: AT&T - IBM COMPATIBLE
16 OPERATING SYSTEM: MS-DOS Version 6.2
17 SOFTWARE: ASCII
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER: US/08/856,841
20 FILING DATE:
21 CLASSIFICATION:
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER: US/08/214,658
24 FILING DATE: 16-MARCH-1994
25 APPLICATION NUMBER: US 07/917,096
26 FILING DATE: 4-AUGUST-1992
27 APPLICATION NUMBER: PCT/DE91/00106
28 FILING DATE: 8-FEBRUARY-1991
29 APPLICATION NUMBER: DE40038262
30 FILING DATE: 8-FEBRUARY-1990
31 ATTORNEY/AGENT INFORMATION:
32 NAME: ROBINSON, WILLIAM R.
33 REGISTRATION NUMBER: 27,224
34 REFERENCE/DOCKET NUMBER: LKR-9222-A
35 TELECOMMUNICATION INFORMATION:
36 TELEPHONE: (212) 697-3355
37 TELEFAX: (212) 557-5635
38 TELEX: NONE
39 INFORMATION FOR SEQ ID NO: 21:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 398
42 TYPE: AMINO ACID
43 TOPOLOGY: LINEAR
44 MOLECULE TYPE:
45 DESCRIPTION: PEPTIDE
46 HYPOTHETICAL: N/A
47 ANTI-SENSE: N/A
48 FRAGMENT TYPE: INTERNAL
49 ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
50 IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
51 POSITION IN GENOME: N/A
52 FEATURE:
53 NAME/KEY: N/A
54 LOCATION: N/A
55 IDENTIFICATION METHOD: amino acid analysis and
56 IDENTIFICATION METHOD: mass spectrometry
57 OTHER INFORMATION:
58 PUBLICATION INFORMATION:
59 AUTHORS: COSSART, Y.E.
60 AUTHORS: FIELD, A.M.
61 AUTHORS: CANT, B.
62 AUTHORS: WIDDOWS, D.
63 TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
64 JOURNAL: LANCET
65 VOLUME: 1
66 ISSUE:
67 PAGES: 72 - 73
68 DATE: 1975
69 DOCUMENT NUMBER:
70 FILING DATE:
71 PUBLICATION DATE:
72 RELEVANT RESIDUES IN SEQ ID NO: 21:
73 US-08-856-841-21

Query Match 100.0%; Score 55; DB 4; Length 398;
Best Local Similarity 100.0%; Pred. No. 0.0038;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKGTQOYTDQ 10
|||||
Db 245 NKGTQOYTDQ 254

RESULT 3

US-08-856-841-20
; Sequence 20, Application US/08856841
; Patent No. 6274307
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHER
; APPLICANT: MANFRED MOTZ
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
; STREET: 99 PARK AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" FLOPPY DISC
; COMPUTER: AT&T - IBM COMPATIBLE
; OPERATING SYSTEM: MS-DOS Version 6.2
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856, 841
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/214, 658
; FILING DATE: 16-MARCH-1994
; APPLICATION NUMBER: US 07/917, 096
; FILING DATE: 4-AUGUST-1992
; APPLICATION NUMBER: PCT/DE91/00106
; FILING DATE: 8-FEBRUARY-1991
; APPLICATION NUMBER: DE40038262
; FILING DATE: 8-FEBRUARY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINSON, WILLIAM R.
; REGISTRATION NUMBER: 27,224
; REFERENCE/DOCKET NUMBER: LKR-9222-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-3355
; TELEFAX: (212) 557-5635
; TELEX: NONE
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: PEPTIDE
; HYPOTHETICAL: N/A
; ANTI-SENSE: N/A
; FRAGMENT TYPE: INTERNAL
; ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
; ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONUM)
; IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
; POSITION IN GENOME: N/A
; FEATURE:
; NAME/KEY: N/A
; LOCATION: N/A
; IDENTIFICATION METHOD: amino acid analysis and
; IDENTIFICATION METHOD: mass spectrometry
; OTHER INFORMATION:

PUBLICATION INFORMATION:

AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO: 20:
US-08-856-841-20

Query Match 100.0%; Score 55; DB 4; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKGTQOYTDQ 10
|||||
Db 256 NKGTQOYTDQ 265

RESULT 4

US-08-856-841-22
; Sequence 22, Application US/08856841
; Patent No. 6274307
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHER
; APPLICANT: MANFRED MOTZ
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
; STREET: 99 PARK AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" FLOPPY DISC
; COMPUTER: AT&T - IBM COMPATIBLE
; OPERATING SYSTEM: MS-DOS Version 6.2
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856, 841
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/214, 658
; FILING DATE: 16-MARCH-1994
; APPLICATION NUMBER: US 07/917, 096
; FILING DATE: 4-AUGUST-1992
; APPLICATION NUMBER: PCT/DE91/00106
; FILING DATE: 8-FEBRUARY-1991
; APPLICATION NUMBER: DE40038262
; FILING DATE: 8-FEBRUARY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINSON, WILLIAM R.
; REGISTRATION NUMBER: 27,224
; REFERENCE/DOCKET NUMBER: LKR-9222-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-3355
; TELEFAX: (212) 557-5635
; TELEX: NONE
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 543
; TYPE: AMINO ACID

TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 22:
US-08-856-841-22

Query Match 100.0%; Score 55; DB 4; Length 543;
Best Local Similarity 100.0%; Pred. No. 0.0052;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NKGTQOYTDO 10
Db 404 NKGTQOYTDO 413

RESULT 5
US-09-554-999-2
Sequence 2, Application US/09554999
Patent No. 6465186
GENERAL INFORMATION:
APPLICANT: Estell, David A.
TITLE OF INVENTION: Proteases from Gram Positive Organisms
FILE REFERENCE: GC389-US
CURRENT APPLICATION NUMBER: US/09/554,999
CURRENT FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: PCT/US98/27040
PRIOR FILING DATE: 1998-12-17
PRIOR APPLICATION NUMBER: GB 9727471.6
PRIOR FILING DATE: 1997-12-30
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 415
TYPE: PRT
ORGANISM: Bacillus subtilis
US-09-554-999-2

Query Match 63.6%; Score 35; DB 4; Length 415;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOYTDO 10
Db 208 NKGTQOYTDO 217

RESULT 6
US-09-095-855-203
Sequence 203, Application US/09095855
Patent No. 6160093
GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: Compounds and Methods for
NUMBER OF SEQUENCES: 208
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/095,855
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/705,347
FILING DATE: 29-AUG-1996
APPLICATION NUMBER: 08/873,970
FILING DATE: 12-JUN-1997
APPLICATION NUMBER: 08/997,362
FILING DATE: 23-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 203:
SEQUENCE CHARACTERISTICS:
LENGTH: 187 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-095-855-203

Query Match 61.8%; Score 34; DB 4; Length 187;
Best Local Similarity 60.0%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOYTDO 10
Db 161 DKSTHTYTNQ 170

RESULT 7
US-09-205-426-203
Sequence 203, Application US/09205426
Patent No. 6406704
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Tan, Paul L. J.
TITLE OF INVENTION: Compounds and Methods for Treatment and
TITLE OF INVENTION: Diagnosis of Mycobacterial Infections
FILE REFERENCE: 11000.1002c4
CURRENT APPLICATION NUMBER: US/09/205,426
CURRENT FILING DATE: 1998-12-04

EARLIER APPLICATION NUMBER: 09/095,855
EARLIER FILING DATE: 1998-06-11
EARLIER APPLICATION NUMBER: 08/997,362
EARLIER FILING DATE: 1997-12-23
EARLIER APPLICATION NUMBER: 08/873,970
EARLIER FILING DATE: 1997-06-12
EARLIER APPLICATION NUMBER: 08/705,347
EARLIER FILING DATE: 1996-08-29
NUMBER OF SEQ ID NOS: 208
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 203
LENGTH: 187
TYPE: PRT
ORGANISM: Mycobacterium vaccae
FEATURE:
NAME/KEY: UNSURE
LOCATION: (186)...(186)
US-09-205-426-203

Query Match 61.8%; Score 34; DB 4; Length 187;
Best local Similarity 60.0%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOXYTDQ 10
: | | | | :
Db 161 DKSTHOXTNQ 170

RESULT 8
US-09-039-198A-14
Sequence 14, Application US/09039198A
Patent No. 6200951
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
APPLICANT: Tjoelker, Larry W.
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,198A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rln-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/34391
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-039-198A-14

Query Match 61.8%; Score 34; DB 4; Length 373;
Best local Similarity 66.7%; Pred. No. 44;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOXYTD 9
: | | | | :
Db 79 NFGTOKFTD 87

RESULT 9
US-09-039-198A-15
Sequence 15, Application US/09039198A
Patent No. 6200951
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
APPLICANT: Tjoelker, Larry W.
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,198A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rln-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/34391
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-039-198A-15

Query Match 61.8%; Score 34; DB 4; Length 373;
Best local Similarity 66.7%; Pred. No. 44;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOXYTD 9
: | | | | :
Db 79 NFGTOKFTD 87

RESULT 10
US-08-877-599-14
Sequence 14, Application US/08877599
Patent No. 6372212
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: Chitinase Materials and Methods
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	61.8%	Score 34	DB 4	Length 373
Best Local Similarity	66.7%	Pred. NO. 44		
Matches	6	Conservative	2	Mismatches
			Indels	Gaps
			1	0

RESULT 11
 US-08-877-599-15
 Sequence 15, Application US/08877599
 Patent No. 6372212
 GENERAL INFORMATION:
 APPLICANT: Gray, Patrick W.
 TITLE OF INVENTION: Chitinase Materials and Methods
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 STREET: 6300 Sears Tower, 233 South Wacker Drive
 City: Chicago
 STATE: Illinois
 COUNTRY: United States of America
 Zip: 60606-6402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/877,599
 FILING DATE:
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/663,618
 FILING DATE: 14-JUN-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Rin-Laures, Li-Hsien
 REGISTRATION NUMBER: 33,547
 REFERENCE/DOCKET NUMBER: 27866/33994
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/474-6300
 TELEFAX: 312/474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 373 amino acids

```

;      TYPE: amino acid
;      STRANDEDNESS: single
;      TOPOLOGY: linear
;      MOLECULE TYPE: peptide
;
US-08-877-599-15

```

Query Match	61.8%	Score 34	DB 4	Length 373
Best Local Similarity	66.7%	Pred. No. 44		
Matches 6, Conservative	2	Mismatches 1	Indels 0	Gaps 0

```

RESULT 12
US-09-267-574-14
: Sequence 14, Application US/09267574
: Patent No. 6395571
: GENERAL INFORMATION:
: APPLICANT: Gray, Patrick W.
: APPLICANT: Tjoelker, Larry W.
: TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
: FILE REFERENCE: 27866/35407
: CURRENT APPLICATION NUMBER: US/09/267,574
: CURRENT FILING DATE: 1999-03-12
: EARLIER APPLICATION NUMBER: 09/039,198
: EARLIER FILING DATE: 1998-03-12
: NUMBER OF SEQ. ID NOS.: 39
: SOFTWARE: patentln Ver. 2.0
: SEQ. ID NO 14
: LENGTH: 373
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-267-574-14

```

Query Match	61.8%	Score 34	DB 4	Length 373
Best Local Similarity	66.7%	Pred. No. 44		
Matches	6	Mismatches	1	Gaps 0
	Conservative			

```

RESULT 13
US-09-267-574-15
: Sequence 15, Application US/09267574
: Patent No. 6398571
: GENERAL INFORMATION:
: APPLICANT: Gray, Patrick W.
: APPLICANT: Joelzer, Larry W.
: TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
: FILE REFERENCE: 27866/35407
: CURRENT APPLICATION NUMBER: US/09/267,574
: EARLIER FILING DATE: 1999-03-12
: EARLIER APPLICATION NUMBER: 09/039,198
: EARLIER FILING DATE: 1998-03-12
: NUMBER OF SEQ. ID NOS: 39
: SOFTWARE: Patentln Ver. 2.0
: SEQ. ID NO. 15
: LENGTH: 373
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-267-574-15

```

Query Match	61.88;	Score 34;	DB 4;	Length 373;
Best Local Similarity	66.7%;	Pred. No. 44;		
Matches	6;	Conservative	2;	Mismatches
			1;	Indels
				Gaps
				0.
QY	1	NGGTQGYTD	9	
	1	:		
b	79	NGGTQGYTD	87	

```

RESULT 14
US-08-486-839-6
; Sequence 6, Application US/08486839
; Patent No. 5928928
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07 - June - 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; US-08-486-839-6

Query Match          61.8%; Score 34; DB 2; Length 387;
Best Local Similarity 66.7%; Pred. No. 46;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOYTD 9
| |||::||
Db 100 NFGTQKFTD 108

RESULT 15
US-09-151-011-6
; Sequence 6, Application US/09151011
; Patent No. 6057142
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A Human Chitinase, Its Recombinant
; TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use In
; TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 6900 Jericho Turnpike
; CITY: Syosset
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11791
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151,011
; FILING DATE: 10 - September - 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Morris, Robert C.
; REGISTRATION NUMBER: 42,910
; REFERENCE/DOCKET NUMBER: 294-32 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; US-09-151-011-6

Query Match          61.8%; Score 34; DB 3; Length 387;
Best Local Similarity 66.7%; Pred. No. 46;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOYTD 9
| |||::||
Db 100 NFGTQKFTD 108

```

Search completed: December 11, 2002, 15:30:21
 Job time : 16 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 15:26:32 : Search time 35 Seconds
(without alignments)
38.072 Million cell updates/sec

Title: US-09-991-433-5
Perfect score: 55
Sequence: 1 NKG2GQYTDQ 10

Scoring table:
BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	370	AA13406	Parvo virus B19 PA
2	55	100.0	543	AA13405	Human parvovirus V
3	55	100.0	554	AAW08987	Erythrovirus V9 VP
4	55	100.0	554	AA123230	Capsid protein enc
5	55	100.0	756	AA171231	Human parvovirus V
6	55	100.0	781	AAW08986	Erythrovirus V9 VP
7	55	100.0	781	AA123227	Drosophila melanog
8	41	74.5	237	ABB60093	Drosophila melanog
9	39	70.9	277	ABB64125	Lactococcus lactis
10	35	63.6	169	ABB54752	Lactococcus lactis

11	35	63.6	415	20	AAV06398	Bacillus subtilis
12	35	63.6	725	23	ABW7613	AMEPY NTPase (AMV)
13	34	61.8	187	20	AAV14929	Amino acid sequenc
14	34	61.8	373	19	AAW40262	Human chitinase pr
15	34	61.8	373	19	AAW40261	Human chitinase pr
16	34	61.8	373	20	AAV42427	Clone of the C-ter
17	34	61.8	373	20	AAV42428	Chitinase amino ac
18	34	61.8	373	23	ABW76293	Human chitinase tr
19	34	61.8	373	23	ABW76294	Human chitinase an
20	34	61.8	387	18	AAW08585	Human 39 kDa chiti
21	34	61.8	466	18	AAW08584	Human 50 kDa chiti
22	34	61.8	466	19	AAW40259	Human chitinase pr
23	34	61.8	466	20	AAV42425	MO-218 clone of hu
24	34	61.8	466	22	AAE00432	Human chitinase pr
25	34	61.8	466	23	ABW76291	Human chitinase..
26	34	61.8	466	23	ABW92136	Herbicidally activ
27	34	61.8	493	20	AAV23206	Sequence encoded b
28	34	61.8	572	13	AAW26574	BrPGS1208 protoxin
29	34	61.8	651	11	AAW06460	Bt isolate 43F. B
30	34	61.8	651	14	AAW33769	Antiscarab pest to
31	34	61.8	651	17	AAW06419	Amino acid sequenc
32	34	61.8	651	20	AAV23197	Amino acid sequenc
33	34	61.8	651	20	AAV23199	Amino acid sequenc
34	34	61.8	651	20	AAV23200	B. thuringiensis Cr
35	34	61.8	651	20	AAV23201	Amino acid sequenc
36	34	61.8	652	12	AAV23202	Amino acid sequenc
37	34	61.8	652	20	AAV23203	Amino acid sequenc
38	34	61.8	652	20	AAV23204	Amino acid sequenc
39	34	61.8	652	20	AAV23205	Amino acid sequenc
40	34	61.8	652	20	AAV23207	Amino acid sequenc
41	34	61.8	652	20	AAV23209	Amino acid sequenc
42	34	61.8	652	20	AAV23174	Amino acid sequenc
43	34	61.8	652	20	AAV23185	Amino acid sequenc
44	34	61.8	652	20	AAV23186	Amino acid sequenc
45	34	61.8	652	20	AAV23186	Amino acid sequenc

ALIGNMENTS

RESULT 1
AA13406 standard; Protein: 370 AA.
ID AA13406;
XX
AC AA13406;
XX
DT 24-OCT-1991 (first entry)
XX
DE Parvo virus B19 PANSE.
XX
KW Primer: PCR; PABST; globulin.
XX
OS Synthetic.
XX
PN DE4003826-A.
XX
PD 14-AUG-1991.
XX
PE 08-FEB-1990; 90DE-4003826.
XX
PR 08-FEB-1990; 90DE-4003826.
XX
PA (MIKR-) MIKROGEN MOLEKULARB.
XX
PI Soutschek E, Motz M;
XX
DR WPI, 1991-246423/34.
XX
PT Immunologically active parvo virus B19 peptide(s) - comprising
PT capsid protein VP1 or VP2 fragments, useful for antibody
PT detection or vaccination
XX
PS Claim 16; Page 10-11; 22pp; German.

```

XX CC VP2 (AAR13405) and its fragments PANSE and PAPST (AAR13407) are useful
CC as immunoassay reagents for detection of anti-B19 antibodies, e.g.
CC for diagnosis of B19 infections, determining the immune status of
CC pregnant woman, testing stored blood, or selecting positive donors
CC for prodn. of B19 hyperimmune globulin preps.
CC The VP2 fragments are expressed by plasmid pUC12VP1, which contains
CC DNA sequences generated from viral DNA by PCR using the primers
CC represented in AAQ13159-63 in pairs.
CC See also AAR13400-07 and AAR13414.
XX
SQ Sequence 370 AA:

Query Match 100.0%; Score 55; DB 12; Length 370;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYTDQ 10
   |||||
   231 NKGTOOYTDQ 240

RESULT 2
AAR13405
ID AAR13405 standard; Protein; 543 AA.
XX
AC AAR13405;
XX
BT 24-OCT-1991 (first entry)
XX
DE Parvo virus B19 VP2.
XX
KM Primer; PCR; globulin; PANSE; PAPST.
XX
OS Synthetic.
XX
PN DE4003826-A.
XX
PD 14-AUG-1991.
XX
PF 08-FEB-1990; 90DE-4003826.
XX
PR 08-FEB-1990; 90DE-4003826.
XX
PA (MIKR-) MIKROGEN MOLEKULARB.
XX
PI Scoutschek E. Motz M;
XX
WPI; 1991-246423/34.

PT Immunologically active parvo virus B19 peptide(s) - comprising
PT capsid protein VP1 or VP2 fragments, useful for antibody
PT detection or vaccination
XX
BT Disclosure; Fig 2-6; 22pp; German.
XX
CC VP2 and its fragments PANSE (AAR13406) and PAPST (AAR13407) are useful
CC as immunoassay reagents for detection of anti-B19 antibodies, e.g.
CC for diagnosis of B19 infections, determining the immune status of
CC pregnant woman, testing stored blood, or selecting positive donors
CC for prodn. of B19 hyperimmune globulin preps.
CC The VP2 fragments are expressed by plasmid pUC12VP1, which contains
CC DNA sequences generated from viral DNA by PCR using the primers
CC represented in AAQ13159-63 in pairs.
CC See also AAR13400-07 and AAR13414.
XX
SQ Sequence 543 AA:

Query Match 100.0%; Score 55; DB 12; Length 543;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYTDQ 10

```

```

Db 404 NKGTOOYTDQ 413
   |||||
   395 NKGTOOYTDQ 404

RESULT 3
AAW08987
ID AAW08987 standard; Protein; 554 AA.
XX
AC AAW08987;
XX
BT 27-FEB-1997 (first entry)
XX
DE Human parvovirus VP-2 protein.
XX
KM Human; parvovirus genome; structural gene; VP-1; VP2; arthritis;
KM non-structural protein; NS; diagnosis; vaccine; parvoviral disease;
KM erythblastemia; abortion; universal fetal hydrops; liver disease;
KM haemorrhagic fever; rheumatism; detection; IgG antibody.
XX
OS Human parvovirus.
XX
PN JP07147986-A.
XX
PD 13-JUN-1995.
XX
PF 24-SEP-1992; 92JP-0281017.
XX
PR 24-SEP-1992; 92JP-0281017.
XX
PA (DENK-) DENKA SEIKEN KK.
PA (ELED ) DENKI KAGAKU KOGYO KK.
XX
DR WPI; 1995-242756/32.
XX
N-PSDB; AAT49535.

PT Human parvovirus gene coding for a polypeptide - useful for
PT developing vaccines against parvoviral diseases such as
PT erythroblastemia, haemorrhagic fever, etc.
XX
PS Claim 3; Page 7-9; 38pp; English.
XX
CC The sequences given in AAW08986 represent the parvovirus structural
CC proteins, VP-1 and VP2, and the non-structural protein, NS. The
CC genomic fragment encoding these proteins may be used for the
CC diagnosis and development of vaccines for parvoviral diseases including
CC erythblastemia, abortion, universal fetal hydrops, liver diseases,
CC haemorrhagic fever, arthritis and rheumatism. The VP-1 and VP-2
CC proteins may be used to detect parvovirus IgG antibodies.
XX
SQ Sequence 554 AA:

Query Match 100.0%; Score 55; DB 16; Length 554;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYTDQ 10
   |||||
   395 NKGTOOYTDQ 404

RESULT 4
AAV23230
ID AAV23230 standard; Protein; 554 AA.
XX
AC AAV23230;
XX
BT 26-AUG-1999 (first entry)
XX
DE Erythrovirus V9 VP2 protein.
XX
KM Erythrovirus V9; differential diagnosis; parvovirus; infection;
KM erythrovirus screening; typing; immunoassay; VP2 protein.
XX

```

OS Erythrovirus.
 XX
 PN FR2771751-A1.
 XX
 PD 04-JUN-1999.
 XX
 PF 03-DEC-1997; 97FR-0015197.
 XX
 PR 03-DEC-1997; 97FR-0015197.
 XX
 PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.
 XX
 PI Auguste V, Garbarg CA, Nguyen QT;
 XX
 DR WPI: 1999-349543/30.
 DR N-PSDB: AAX81586.
 XX
 PT Erythrovirus V9 and its nucleic acid sequences - can be used in the
 PT diagnosis of its infections
 XX
 SS Claim 19; Page 57-58; 80pp; French.
 XX
 CC The present sequence represents an erythrovirus V9 protein.
 CC Probes and primers derived from erythrovirus V9 polynucleotide
 CC sequences (AAX81580) can be used for differential diagnosis of
 CC erythrovirus (parvovirus) infections by a combination of
 CC amplification and hybridisation assay. The probes can also be
 CC used to assess susceptibility to erythrovirus infection and
 CC for erythrovirus screening and typing. The antibodies can be
 CC used in immunoassays for diagnosis of erythrovirus V9 infections.
 XX
 SQ Sequence 554 AA;
 XX
 Query Match 100.0%; Score 55; DB 20; Length 554;
 Best Local Similarity 100.0%; Pred. No. 0.028;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 NKGTOQYTDQ 10
 Db 395 NKGTOQYTDQ 404
 XX
 RESULT 5
 AAY71231
 ID AAY71231 standard; Protein: 756 AA.
 XX
 AC AAY71231;
 XX
 DE 08-SEP-2000 (first entry)
 XX
 DE Capsid protein encoded by AAV2/human parvovirus B19 chimeric vector.
 XX
 KM Adeno-associated virus; AAV2; human parvovirus B19; chimeric;
 KM recombinant parvoviral vector; cellular tropism; cap protein;
 KM capsid; gene delivery; gene therapy; VP1; VP2; VP3.
 XX
 OS Chimeric - Adeno associated virus serotype 2.
 OS Chimeric - Human parvovirus B19.
 XX
 PN WO200028004-A1.
 XX
 PD 18-MAY-2000.
 XX
 PF 10-NOV-1999; 99WO-US26505.
 XX
 PR 10-NOV-1998; 98US-0107840.
 PR 10-MAR-1999; 99US-0123651.
 XX
 PA (UYNC-) UNIV NORTH CAROLINA.
 XX
 PI Rabinowitz JE, Samulski RJ, Xiao W;
 XX
 DR WPI: 2000-376523/32.

DR N-PSDB; AAD00833.
 XX
 PT Recombinant parvoviral vectors with altered packaging, tropisms and
 PT immunogenic properties, useful in gene therapy protocols -
 XX
 PS Example 21; Page 142; 153pp; English.
 XX
 CC The patent discloses modified parvovirus vectors with advantageous
 CC antigenic properties, packaging capabilities and cellular tropisms.
 CC These vectors can be used in standard recombinant DNA protocols e.g. gene
 CC therapy for delivering nucleic acids to cells.
 CC The present sequence is a capsid protein encoded by an
 CC adeno-associated virus serotype 2 (AAV2)/human parvovirus B19 chimeric
 CC vector. This vector encodes AAV2 VP1 and VP2 capsid proteins
 CC and human parvovirus B19 VP2 protein. The chimeric vector was
 CC constructed by replacing the VP3 major cap protein of AAV2 with B19's
 CC VP2. Recombinant parvovirus comprising the chimeric capsid is useful
 CC for gene delivery.
 XX
 SQ Sequence 756 AA;
 XX
 Query Match 100.0%; Score 55; DB 21; Length 756;
 Best Local Similarity 100.0%; Pred. No. 0.04;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 NKGTOQYTDQ 10
 Db 597 NKGTOQYTDQ 606
 XX
 RESULT 6
 AAW08986
 ID AAW08986 standard; Protein: 781 AA.
 XX
 AC AAW08986;
 XX
 DE 27-FEB-1997 (first entry)
 XX
 DE Human parvovirus VP-1 protein.
 XX
 KM Human; parvovirus genome; structural gene; VP-1; VP2; arthritis;
 KM non-structural protein; NS; diagnosis; vaccine; parvoviral disease;
 KM erythblastemia; abortion; universal fetal hydrops; liver disease;
 KM haemorrhagic fever; rheumatism; detection; IgG antibody.
 XX
 OS Human parvovirus.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 19
 FT /label= Gly, Val
 FT Misc-difference 61
 FT /label= Asn, Asp
 FT Misc-difference 220
 FT /label= His, Asn
 XX
 PN JP07147986-A.
 XX
 PD 13-JUN-1995.
 XX
 PF 24-SEP-1992; 92JP-0281017.
 XX
 PR 24-SEP-1992; 92JP-0281017.
 XX
 PA (DENK-) DENKA SEIKEN KK.
 PA (ELED) DENKI KAGAKU KOGYO KK.
 XX
 DR WPI: 1995-242756/32.
 DR N-PSDB: AAT49535.
 XX
 PT Human parvovirus gene coding for a polypeptide - useful for
 PT developing vaccines against parvoviral diseases such as
 PT erythroblastemia, haemorrhagic fever, etc.
 XX

PS Claim 2; Page 5-7; 38pp; English.
XX
XX The sequences given in AAM08986 represent the parvovirus structural
CC proteins, VP-1 and VP-2, and the non-structural protein, NS. The
CC genomic fragment encoding these proteins may be used for the
CC diagnosis and development of vaccines for parvoviral diseases including
CC erythblastemia, abortion, universal fetal hydrops, liver diseases,
CC haemorrhagic fever, arthritis and rheumatism. The VP-1 and VP-2
CC proteins may be used to detect parvovirus IgG antibodies.
XX
SQ Sequence 781 AA:
Query Match 100.0%; Score 55; DB 16; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NKGTOQYTDQ 10
Db 622 NKGTOQYTDQ 631

RESULT 7
ID AAY23227 standard; Protein: 781 AA.
XX
XX AAY23227;
AC
XX
XX 26-AUG-1999 (first entry)
DT
XX
DE Erythrovirus V9 VP1 protein.
XX
XX Erythrovirus V9; differential diagnosis: parvovirus; infection;
KM erythrovirus screening; typing; immunosassay; VP1 protein.
XX
OS Erythrovirus.
XX
XX FR2771751-A1.
PN
XX
XX 04-JUN-1999.
PD
XX
XX 03-DEC-1997; 97FR-0015197.
PF
XX
XX 03-DEC-1997; 97FR-0015197.
PR
XX
XX (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.
PA
XX
PI Auguste V, Garbary CA, Nguyen QT;
WPI: 1999-349543/30.
N-PSDB; AAX81583.
XX
PT Erythrovirus V9 and its nucleic acid sequences - can be used in the
XX diagnosis of its infections
XX
XX
PS Claim 19; Page 50-52; 80pp; French.
XX
XX The present sequence represents an erythrovirus V9 protein.
CC Probes and primers derived from erythrovirus V9 polynucleotide
CC sequences (AAX81580) can be used for differential diagnosis of
CC erythrovirus (parvovirus) infections by a combination of
CC amplification and hybridisation assay. The probes can also be
CC used to assess susceptibility to erythrovirus infection and
CC for erythrovirus screening and typing. The antibodies can be
CC used in immunosassays for diagnosis of erythrovirus V9 infections.
XX
SQ Sequence 781 AA:
Query Match 100.0%; Score 55; DB 20; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NKGTOQYTDQ 10
XX
XX *****

Db 622 NKGTOQYTDQ 631

RESULT 8
ID ABB60093 standard; Protein: 237 AA.
XX
XX ABB60093;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 7071.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
XX Drosophila melanogaster.
OS
XX
XX WO200171042-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US09231.
PF
XX
XX 23-MAR-2000; 2000US-191637P.
PR
XX
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX
XX WPI: 2001-656860/75.
DR
XX
XX N-PSDB; ABL04196.
DR
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
PT
XX
XX Disclosure; SEQ ID NO 7071; 21pp + Sequence Listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABJ30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 237 AA:
Query Match 74.5%; Score 41; DB 22; Length 237;
Best Local Similarity 70.0%; Pred. No. 5.7;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 NKGTOQYTDQ 10
Db 30 NNGTOKYTNQ 39

RESULT 9
ID ABB64125 standard; Protein: 277 AA.
XX
XX ABB64125;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 19167.
DE
XX

KW	Drosophilid; developmental biology; cell signalling; insecticide;
KW	pharmaceutical.
XX	
OS	Drosophila melanogaster.
PN	WO200171042-A2.
XX	
PD	27-SEP-2001.
XX	
PF	23-MAR-2001; 2001WO-US09231.
XX	
PR	23-MAR-2000; 2000US-191637P.
PR	11-JUL-2000; 2000US-0614150.
PA	(PEXE) PE CORP NY.
PI	Venter JC, Adams M, Li PWD, Myers EW;
PI	
DR	WPI: 2001-656860/75.
DR	N-PSDB; ABL08228.
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more
PT	genes from Drosophila and for elucidating cell signalling and cell-cell
PT	interactions -
XX	
PS	Disclosure: SEQ ID NO 19167; 21pp + Sequence Listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent
CC	capable of detecting 1000 or more genes from Drosophila. The invention is
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC	sequences (AB101840-AB116175) and the encoded proteins
CC	(AB157737-AB172072).
CC	The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	atftp.wipo.int/pub/published_pct_sequences.
XX	
SO	Sequence 277 AA;
XX	
Query Match	70.9%; Score 39; DB 22; Length 277;
Best Local Similarity	60.0%; Pred. No. 17;
Matches 6; Conservative	3; Mismatches 1; Indels 0; Gaps 0.
Qy	1 NKGTOOYTDQ 10
	1:11:111:
Db	247 NNGTROYEDE 256
XX	
RESULT 10	
ID	ABB54752
XX	ABB54752 standard; Protein; 169 AA.
XX	
AC	ABB54752;
XX	
DT	I6-MAY-2002 (first entry)
XX	
DE	Lactococcus lactis protein p1339.
XX	
KW	Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
XX	
OS	Lactococcus lactis IL1403.
XX	
PN	FR2807446-A1.
XX	
PD	12-OCT-2001.
XX	
PF	11-APR-2000; 2000FR-0004630.
XX	
PR	11-APR-2000; 2000FR-0004630.
XX	
PA	(INRG) INRA INST NAT RECH AGRONOMIQUE.

```

XX Bolotine A, Sorokine A, Renault P, Enrich SD;
PI
XX
DR WPI: 2002-043418/06.
PT New nucleotide sequence useful in the identification or Lactococcus
PT lactis and related species -
XX
PS Claim 6; SEQ ID No 1454; 2504pp; French.
CC The present invention is related to a Lactococcus lactis nucleotide
CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The
CC nucleic acid sequence is useful in the detection and/or amplification of
CC nucleic acid sequence, particularly to identify Lactococcus lactis or
CC related species. The proteins of the invention are useful for the
CC biosynthesis or biodegradation of a composition of interest. The
CC invention helps research in lactic bacteria, particularly useful in the
CC production of yogurt and cheese.
CC Note: The sequence data for this patent is based on equivalent patent
CC WO200177334 (published 18-OCT-2001) which is available in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 169 AA;
Query Match 63.6%; Score 35; DB 23; Length 169;
Best Local Similarity 60.0%; Pred. No. 58;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0.
OY 1 NKGTOQYTDQ 10
| :||| |
Db 60 NNSSQYADQ 69
RESULT 11
AY06398
ID AAY06398 standard; Protein; 415 AA.
AC AAY06398;
XX
DT 20-SEP-1999 (first entry)
XX
DE Bacillus subtilis metalloprotease YmfH.
DE
KW Metalloprotease; protease; YmfH; detergent; surfactant; cleaning;
RW textile; feedstuff; animal feed; host cell.
XX
OS Bacillus subtilis.
XX
FH Key Location/Qualifiers
FT Domain 55..59
FT /note= "catalytic domain HXXEH motif"
PN WO9933960-A2.
XX
PD 08-JUL-1999.
XX
PF 17-DEC-1998; 98WO-US27040.
XX
PR 30-DEC-1997; 97GB-0027471.
XX
PA (GENV ) GENENCOR INT INC.
PI Estell DA;
XX
DR WPI: 1999-419100/35.
DR N-PSDB; AAX59344.
XX
A Bacillus subtilis metalloprotease, designated YmfH, useful in
cleaning compositions, animal feed and for treating textiles
Claim 9; Fig 1A-F; 32pp: English.
XX
The present sequence represents a novel metalloprotease (MP),

```

CC designated ymfH, of Bacillus subtilis. ymfH DNA (see AAX59344) was
CC identified via a BLAST search of B. subtilis genomic DNA. The
CC deduced protein sequence shows identity to the MP pili-lysin of
CC Escherichia coli, including the presence of the catalytic domain
CC HXXHR+E. An expression vector including ymfH DNA and a host cell
CC comprising the vector are claimed. Also claimed are a cleaning
CC composition, an animal feed and a composition for the treatment of
CC a textile, all comprising ymfH. Gram positive microorganisms
CC having a mutation or deletion of all or part of ymfH DNA are used
CC as host cells for expression of a homologous or heterologous
CC protein, such as a hormone, growth factor, cytokine or enzyme,
CC especially a protease, carbonhydrase, lipase, isomerase, oxidase,
CC reductase, transferase, kinase or phosphatase (all claimed). Also
CC claimed is a method for detecting a Gram positive microorganism MP
CC using a probe comprising all or part of the ymfH DNA.

XX Sequence 415 AA:

Query Match 63.6%; Score 35; DB 20; Length 415;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOOYTDQ 10
||:|||||

Db 208 NQKKPTDQ 217

RESULT 12
AB77613
ID ABB77613 standard; Protein: 726 AA.
XX ABB77613;
XX
XX 01-JUL-2002 (first entry)
XX
XX AMEPV NTPase (AMV087).
XX
XX AMEPV; gene therapy; viral vector; chromosome mapping; gene mapping;
XX genetic deficiency disorder; NTPase.
XX
XX Amsacta moorei entomopoxvirus.
XX
XX WO200212526-A2.
XX
XX 14-FEB-2002.
XX
XX 10-AUG-2001; 2001MO-US25287.
XX
XX 10-AUG-2000; 2000US-224479P.
XX 14-SEP-2000; 2000US-0662254.
XX (UYFL) UNIV FLORIDA.
XX
XX Moyer RW, Li Y, Bawden AL;
XX
XX WPI: 2002-227161/28.
XX N-PSDB; ABL56216.
XX
XX Novel recombinant entomopox virus vector useful for delivering
XX polynucleotide encoding protein to vertebrate cell, comprises
XX polynucleotide encoding protein operably linked with heterologous
XX promoter sequence -
XX
XX Claim 75; Page 263-266; 326pp; English.

CC The invention relates to a recombinant entomopox virus (EPV) vector,
CC comprising a polynucleotide encoding a protein operably linked with a
CC heterologous promoter sequence. The invention also concerns methods for
CC providing gene therapy for genetic deficiency disorders. Vectors of the
CC invention are useful for delivering a polynucleotide encoding a protein
CC to a vertebrate cell preferably a mammalian cell, such as a human cell.
CC The vector is introduced into the vertebrate cell by infection in a viral
CC particle, or by transfection, transduction, or injection either in vitro

CC or in vivo. The vector is useful for the delivery and expression of
CC biologically useful proteins in gene therapy protocols, and for
CC delivering large DNA segments for engineering of vertebrate cells.
CC Polynucleotides of the invention have applications in techniques such as
CC their use as insertion sites for foreign genes of interest, hybridisation
CC probes, for chromosome and gene mapping, in PCR technologies, and in the
CC production of sense or antisense nucleic acids. Vectors of the invention
CC provide for stable integration and expression of heterologous DNA in host
CC cells, and are adapted for accepting large heterologous polynucleotide
CC inserts which can be delivered in an infected or transformed cell and
CC expressed in a stable fraction. The current sequence represents an
CC amsacta moorei entomopoxvirus (AMEPV) NTPase (AMV087).

XX Sequence 726 AA:

Query Match 63.6%; Score 35; DB 23; Length 726;
Best Local Similarity 60.0%; Pred. No. 3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOOYTDQ 10
||:|||||

Db 598 NKNIKOYTEQ 607

RESULT 13
AA14929
ID AA14929 standard; Protein: 187 AA.
XX AA14929;
XX
XX 25-OCT-1999 (first entry)
XX
XX Amino acid sequence of M. vaccae antigen GV-41B.
XX
XX
XX Mycobacterium vaccae protein; antigen; T cell activation; cytokine;
XX dendritic cell maturation; infectious disease; immune disorder; cancer;
XX respiratory system; mycobacterial infection; allergy; tuberculosis;
XX leprosy; sarcoidosis; lung cancer; asthma; skin disorder; psoriasis;
XX dermatitis; eczema; alopecia areata; skin cancer; basal carcinoma;
XX squamous cell carcinoma; melanoma.
XX
XX Mycobacterium vaccae.
XX
XX WO932634-A2.
XX
XX 01-JUL-1999.
XX
XX 23-DEC-1998; 98MO-NZ00189.
XX
XX 04-DEC-1998; 98US-0205426.
XX 23-DEC-1997; 97US-0996624.
XX 23-DEC-1997; 97US-0997080.
XX 23-DEC-1997; 97US-0997362.
XX 11-JUN-1998; 98US-0095855.
XX 17-SEP-1998; 98US-0156181.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX
XX Prestidge RU, Skinner MA, Tan P, Vlisser ES, Watson J;
XX WPI: 1999-430163/36.
XX N-PSDB; AA211394.
XX
XX Enhancing immune response to an antigen
XX
XX Claim 1; Page 240; 243pp; English.

CC The invention provides heat-killed Mycobacterium vaccae, or recombinant
CC M. vaccae proteins. The M. vaccae proteins may be employed to activate
CC T cells and natural killer cells, to stimulate the production of
CC cytokines, to enhance the expression of co-stimulatory molecules on
CC dendritic cells and monocytes, and to enhance dendritic cell maturation
CC and function. The proteins can be expressed by standard recombinant

CC methodology. Pharmaceutical compositions comprising the proteins or
 CC nucleic acid sequences encoding the proteins can be used for the
 CC treatment, prevention, and detection of disorders including infectious
 CC diseases, immune disorders and cancer. In particular, the compounds and
 CC methods are used for treatment of diseases of the respiratory system,
 CC such as mycobacterial infections, asthma, allergies, tuberculosis,
 CC leprosy, sarcoidosis and lung cancers, and disorders of the skin such as
 CC psoriasis, atopic dermatitis, eczema, allergic contact dermatitis,
 CC alopecia areata, and skin cancers such as basal carcinoma, squamous cell
 CC carcinoma and melanoma.

XX
 XX
 SQ Sequence 187 AA;
 Query Match 61.8%; Score 34; DB 20; Length 187;
 Best Local Similarity 60.0%; Pred. No. 1e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQQTQTDQ 10
 : | | | | | | | | | |
 Db 161 DKSTHOYTNO 170

RESULT 14
 AAW40262
 ID AAW40262 standard; Protein: 373 AA.
 XX
 AC AAW40262;
 XX
 DT 15-JUN-1998 (first entry)
 XX
 DE Human chitinase protein analogue.
 XX
 KW Chitinase: human; fungal infection; immunogen; diagnosis; treatment;
 KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;
 KW rheumatoid arthritis; overexpression; extracellular matrix.
 XX
 OS Homo sapiens.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 370 /label= p370S
 FT /note= "Wild type Pro is replaced by Ser"
 XX
 PN WO9747752-A1.
 XX
 PD 18-DEC-1997.
 XX
 FF 16-JUN-1997; 97WO-US10460.
 XX
 PR 14-JUN-1996; 96US-0663618.
 XX
 PA (ICOS-) ICOS CORP.
 XX
 PI Gray PW;
 XX
 DR WPI; 1998-052316/05.
 XX
 PT Nucleic acids encoding human chitinase - useful as antifungal
 PT agents, especially in combination with other antifungals
 XX
 PS Claim 29; Page 49-50; 63pp; English.

XX This sequence encodes a fragment of a novel human chitinase protein
 CC analogue in which a proline at position 370 of the wild type protein
 CC (see AAW40261) is substituted with a serine and the C-terminal 72
 CC residues of the mature protein are deleted. These residues are not
 CC critical to chitinase enzymatic activity. Chitinases are useful for
 CC treating or preventing fungal infection and as immunogens for generating
 CC antibodies which are used to purify, detect and quantify chitinases, e.g.
 CC for diagnosis of Gaucher's disease. The nucleic acid sequence of the
 CC chitinase is also useful as a probe to identify and isolate genomic DNA
 CC encoding chitinases or similar proteins, or cells expressing them or to

CC generate transgenic ('knockout') rodents. It can also be used in
 CC hybridisation assays and to detect genetic alterations in the chitinase
 CC gene related to disease. Agents that inhibit this protein may be useful
 CC in treatment of Gaucher's disease and rheumatoid arthritis, where
 CC overexpression of the protein can damage the extracellular matrix.
 CC Chitinase also improves the activity of other antifungal agents and may
 CC allow a reduction in the dose of such agents, and thus of their side
 CC effects.

XX
 XX
 SQ Sequence 373 AA;
 Query Match 61.8%; Score 34; DB 19; Length 373;
 Best Local Similarity 66.7%; Pred. No. 2.2e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQQTQTD 9
 : | | | | | | | | | |
 Db 79 NFGTQKFTMD 87

RESULT 15
 AAW40261
 ID AAW40261 standard; Protein: 373 AA.
 XX
 AC AAW40261;
 XX
 DT 15-JUN-1998 (first entry)
 XX
 DE Human chitinase protein fragment.
 XX
 KW Chitinase: human; fungal infection; immunogen; diagnosis; treatment;
 KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;
 KW rheumatoid arthritis; overexpression; extracellular matrix.
 XX
 OS Homo sapiens.
 OS
 PN WO9747752-A1.
 XX
 PD 18-DEC-1997.
 XX
 FF 16-JUN-1997; 97WO-US10460.
 XX
 PR 14-JUN-1996; 96US-0663618.
 XX
 PA (ICOS-) ICOS CORP.
 XX
 PI Gray PW;
 XX
 DR WPI; 1998-052316/05.
 XX
 DR N-PSDB; AAV10435.
 XX
 PT Nucleic acids encoding human chitinase - useful as antifungal
 PT agents, especially in combination with other antifungals
 XX
 PS Claim 23; Page 48-49; 63pp; English.

XX This sequence encodes a fragment of a novel human chitinase protein
 CC which lacks the C-terminal 72 residues of the mature protein. These
 CC residues are not critical to chitinase enzymatic activity. Chitinases are
 CC useful for treating or preventing fungal infection and as immunogens for
 CC generating antibodies which are used to purify, detect and quantify
 CC chitinases, e.g. for diagnosis of Gaucher's disease. The nucleic acid
 CC sequence of the chitinase is also useful as a probe to identify and
 CC isolate genomic DNA encoding chitinases or similar proteins, or cells
 CC expressing them or to generate transgenic ('knockout') rodents. It can
 CC also be used in hybridisation assays and to detect genetic alterations
 CC in the chitinase gene related to disease. Agents that inhibit this
 CC protein may be useful in treatment of Gaucher's disease and rheumatoid
 CC arthritis, where overexpression of the protein can damage the
 CC extracellular matrix. Chitinase also improves the activity of other
 CC antifungal agents and may allow a reduction in the dose of such agents,
 CC and thus of their side effects.

SQ Sequence 373 AA;

Query Match 61.8%; Score 34; DB 19; Length 373;

Best Local Similarity 66.7%; Pred. No. 2.2e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQRYTD 9

| | | | |

Db 79 NFGTQKFTD 87

Search completed: December 11, 2002, 15:28:45
Job time : 36 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 15:25:32 : Search time 14 Seconds
(Without alignments)
42.033 Million cell updates/sec

Title: US-09-991-433-48
Perfect score: 110
Sequence: 1 TYFPNKGTYQYTDQIERPLM 20

Scoring table: BLOSUM62
Gap 10.0 , Gapect 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	395	4	US-08-856-841-13
2	110	100.0	398	4	US-08-856-841-21
3	110	100.0	415	4	US-08-856-841-20
4	110	100.0	543	4	US-08-856-841-22
5	45	40.9	383	4	US-09-134-001C-3701
6	43.5	39.5	763	2	US-08-677-862-2
7	43.5	39.5	763	2	US-09-252-571-2
8	43.5	39.5	763	3	US-09-434-065-2
9	43.5	39.5	763	4	US-08-789-275-4
10	43.5	39.5	763	4	US-08-789-275-5
11	43	39.1	720	1	US-07-731-157A-2
12	43	39.1	720	2	US-08-541-780-2
13	43	38.2	1501	2	US-08-447-464-3
14	42	38.2	1501	1	US-08-716-679-3
15	42	38.2	1911	1	US-08-348-006B-5
16	42	38.2	1911	2	US-08-800-825A-5
17	42	38.2	1911	4	US-09-158-657-5
18	42	38.2	1911	5	PCT-US94-10166-5
19	40	36.4	187	4	US-09-095-855-203
20	40	36.4	187	4	US-09-205-426-203
21	40	36.4	1872	4	US-09-220-641-3
22	40	36.4	1872	1	US-08-188-582-14
23	40	36.4	1872	1	US-08-646-715-14
24	40	36.4	1893	1	US-08-188-582-11
25	40	36.4	1893	1	US-08-646-715-11
26	39.5	35.9	510	1	US-08-249-112-3
27	39.5	35.9	510	5	PCT-US95-06556-3

28	39	35.5	16	3	US-08-931-220-2	Sequence 2, Appl
29	39	35.5	16	5	PCT-US95-11723-2	Sequence 2, Appl
30	39	35.5	16	5	PCT-US96-05997-15	Sequence 15, Appl
31	39	35.5	241	4	US-09-134-001C-5493	Sequence 3493, Ap
32	39	35.5	398	3	US-08-931-220-5	Sequence 5, Appl
33	39	35.5	398	5	PCT-US95-11723-5	Sequence 5, Appl
34	39	35.5	398	5	PCT-US96-05997-1	Sequence 1, Appl
35	39	35.5	622	2	US-08-356-786-16	Sequence 16, Appl
36	38	34.5	118	2	US-08-846-762-10	Sequence 10, Appl
37	38	34.5	133	4	US-09-134-001C-4122	Sequence 1, Appl
38	38	34.5	414	2	US-08-984-171-1	Sequence 4122, Ap
39	38	34.5	430	4	US-09-134-001C-4302	Sequence 4302, Ap
40	38	34.5	1158	4	US-09-060-482-2	Sequence 2, Appl
41	37.5	34.1	108	4	US-09-240-274-51	Sequence 51, Appl
42	37.5	34.1	108	4	US-09-240-274-52	Sequence 52, Appl
43	37.5	34.1	237	6	516941-3	Patent No. 516941
44	37.5	34.1	1076	4	US-09-470-443-6	Sequence 6, Appl
45	37.5	34.1	1145	4	US-09-470-443-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-856-841-13
Sequence 13, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHER
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAYDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856, 841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214, 658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917, 096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27, 224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO.: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 395
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE

```

1  HYPOTHEICAL: N/A
2  ANTI-SENSE: N/A
3  FRAGMENT TYPE: INTERNAL
4  ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
5  ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
6  IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
7  POSITION IN GENOME: N/A
8  FEATURE:
9  NAME/KEY: N/A
10 LOCATION: N/A
11 IDENTIFICATION METHOD: amino acid analysis and
12 IDENTIFICATION METHOD: mass spectrometry
13 OTHER INFORMATION:
14 PUBLICATION INFORMATION:
15 AUTHORS: COSSART, Y. E.
16 AUTHORS: FIELD, A. M.
17 AUTHORS: CANT, B.
18 AUTHORS: WIDOWS, D.
19 TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
20 JOURNAL: LANCET
21 VOLUME: I
22 ISSUE:
23 PAGES: 72 - 73
24 DATE: 1975
25 DOCUMENT NUMBER:
26 FILING DATE:
27 PUBLICATION DATE:
28 RELEVANT RESIDUES IN SEQ ID NO: 13:
29 PUBLICATION INFORMATION:
30 AUTHORS: MANIATIS, T.
31 AUTHORS: FRITSCH, E. F.
32 AUTHORS: SAMBROOK, J.
33 TITLE: MOLECULAR CLONING
34 JOURNAL: COLD SPRING HARBOR, NY
35 VOLUME:
36 ISSUE:
37 PAGES:
38 DATE: 1982
39 DOCUMENT NUMBER:
40 FILING DATE:
41 PUBLICATION DATE:
42 RELEVANT RESIDUES IN SEQ ID NO: 13:
43 PUBLICATION INFORMATION:
44 AUTHORS: SMITH, D. B.
45 AUTHORS: JOHNSON, K. S.
46 TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES
47 TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH
48 TITLE: GLUTATHIONE S. TRANSFERASE
49 JOURNAL: GENE
50 VOLUME:
51 ISSUE:
52 PAGES: 67
53 DATE: 1988
54 DOCUMENT NUMBER:
55 FILING DATE:
56 PUBLICATION DATE:
57 RELEVANT RESIDUES IN SEQ ID NO: 13:
58 US-08-856-841-13
59
60 Query Match 100.0%; Score 110; DB 4; Length 395;
61 Best Local Similarity 100.0%; Pred. No. 1,1e-10;
62 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
63
64 Oy 1 TYFPNKGTOQYTDQIERPLM 20
65 ||||||||||||||||
66 Db 252 TYFPNKGTOQYTDQIERPLM 271
67
68 RESULT 2
69 US-08-856-841-21
70 Sequence 21, Application US/08856841
71 Patent No. 6274307
72 GENERAL INFORMATION:

```

APPLICANT: ERWIN SOUTSCHEK
 APPLICANT: MARNER MORTZ
 TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
 TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
 NUMBER OF SEQUENCES: 28
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: BROOKS HAIDT HAEFFNER & DELAHUNTY
 STREET: 99 PARK AVENUE
 CITY: NEW YORK
 STATE: NY
 COUNTRY: USA
 ZIP: 10016
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" FLOPPY DISC
 COMPUTER: AT&T - IBM COMPATIBLE
 OPERATING SYSTEM: MS-DOS Version 6.2
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/856,841
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/214,658
 FILING DATE: 16-MARCH-1994
 APPLICATION NUMBER: US 07/917,096
 FILING DATE: 4-AUGUST-1992
 APPLICATION NUMBER: PCT/DE91/00106
 FILING DATE: 8-FEBRUARY-1991
 APPLICATION NUMBER: DE40038262
 FILING DATE: 8-FEBRUARY-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: ROBINSON, WILLIAM R.
 REGISTRATION NUMBER: 27,224
 REFERENCE/DOCKET NUMBER: LKR-9222-A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 697-3355
 TELEFAX: (212) 557-5635
 TELEX: NONE
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 398
 TYPE: AMINO ACID
 TOPOLOGY: LINEAR
 MOLECULE TYPE:
 DESCRIPTION: PEPTIDE
 HYPOTHEICAL: N/A
 ANTI-SENSE: N/A
 FRAGMENT TYPE: INTERNAL
 ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
 IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
 POSITION IN GENOME: N/A
 FEATURE:
 NAME/KEY: N/A
 LOCATION: N/A
 IDENTIFICATION METHOD: amino acid analysis and
 IDENTIFICATION METHOD: mass spectrometry
 OTHER INFORMATION:
 PUBLICATION INFORMATION:
 AUTHORS: COSSART, Y.E.
 AUTHORS: FIELD, A.M.
 AUTHORS: CANT, B.
 AUTHORS: WIDOWS, D.
 TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
 JOURNAL: LANCET
 VOLUME: 1
 ISSUE:
 PAGES: 72 - 73
 DATE: 1975
 DOCUMENT NUMBER:
 FILING DATE:
 PUBLICATION DATE:
 RELEVANT RESIDUES IN SEQ ID NO: 21:
 15-08-856-841-21

Query Match 100.0%; Score 110; DB 4; length 398;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYFPNKGTQOYTDQIERPLM 20
|||||
DB 241 TYFPNKGTQOYTDQIERPLM 260

RESULT 3
US-08-856-841-20

Sequence 20, Application US/08856841
Patent No. 6274307

GENERAL INFORMATION:

APPLICANT: ERWIN SOUTSCHEK

TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES

TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY

STREET: 99 PARK AVENUE

CITY: NEW YORK

STATE: NY

COUNTRY: USA

ZIP: 10016

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" FLOPPY DISC

COMPUTER: AT&T - IBM COMPATIBLE

OPERATING SYSTEM: MS-DOS Version 6.2

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/856, 841

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/214, 658

FILING DATE: 16-MARCH-1994

APPLICATION NUMBER: US 07/917, 096

FILING DATE: 4-AUGUST-1992

APPLICATION NUMBER: PCT/DE91/00106

FILING DATE: 8-FEBRUARY-1991

APPLICATION NUMBER: DE40038262

FILING DATE: 8-FEBRUARY-1990

ATTORNEY/AGENT INFORMATION:

NAME: ROBINSON, WILLIAM R.

REGISTRATION NUMBER: 27,224

REFERENCE/DOCKET NUMBER: LKR-9222-A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 697-3355

TELEFAX: (212) 557-5635

TELEX: NONE

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 415

TYPE: AMINO ACID

TOPOLOGY: LINEAR

MOLECULE TYPE:

DESCRIPTION: PEPTIDE

HYPOTHETICAL: N/A

ANTI-SENSE: N/A

FRAGMENT TYPE: INTERNAL

ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE

ORIGINAL SOURCE: INFECTION (ERTHEMA INFECTIOSUM)

IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE

POSITION IN GENOME: N/A

PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: I
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 20:

US-08-856-841-20

Query Match 100.0%; Score 110; DB 4; length 415;
Best Local Similarity 100.0%; Pred. No. 1.2e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYFPNKGTQOYTDQIERPLM 20
|||||
DB 252 TYFPNKGTQOYTDQIERPLM 271

RESULT 4
US-08-856-841-22

Sequence 22, Application US/08856841
Patent No. 6274307

GENERAL INFORMATION:

APPLICANT: ERWIN SOUTSCHEK

TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES

TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY

STREET: 99 PARK AVENUE

CITY: NEW YORK

STATE: NY

COUNTRY: USA

ZIP: 10016

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" FLOPPY DISC

COMPUTER: AT&T - IBM COMPATIBLE

OPERATING SYSTEM: MS-DOS Version 6.2

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/856, 841

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/214, 658

FILING DATE: 16-MARCH-1994

APPLICATION NUMBER: US 07/917, 096

FILING DATE: 4-AUGUST-1992

APPLICATION NUMBER: PCT/DE91/00106

FILING DATE: 8-FEBRUARY-1991

APPLICATION NUMBER: DE40038262

FILING DATE: 8-FEBRUARY-1990

ATTORNEY/AGENT INFORMATION:

NAME: ROBINSON, WILLIAM R.

REGISTRATION NUMBER: 27,224

REFERENCE/DOCKET NUMBER: LKR-9222-A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 697-3355

TELEFAX: (212) 557-5635

TELEX: NONE

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 543

TYPE: AMINO ACID

```

? TOPOLOGY: LINEAR
? MOLECULE TYPE:
? DESCRIPTION: PEPTIDE
? HYPOTHETICAL: N/A
? ANTI-SENSE: N/A
? FRAGMENT TYPE: INTERNAL
? ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
? ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
? IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
? POSITION IN GENOME: N/A
? FEATURE:
? NAME/KEY: N/A
? LOCATION: N/A
? IDENTIFICATION METHOD: amino acid analysis and
? IDENTIFICATION METHOD: mass spectrometry
? OTHER INFORMATION:
? PUBLICATION INFORMATION:
? AUTHORS: COSSART, Y.E.
? AUTHORS: FIELD, A.M.
? AUTHORS: CANT, B.
? AUTHORS: WIDDOWS, D.
? TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
? JOURNAL: LANCET
? VOLUME: 1
? ISSUE:
? PAGES: 72 - 73
? DATE: 1975
? DOCUMENT NUMBER:
? FILING DATE:
? PUBLICATION DATE:
? RELEVANT RESIDUES IN SEQ ID NO: 22:
US-08-856-841-22

```

```

Query Match 100.0%; Score 110; DB 4; Length 543;
Best Local Similarity 100.0%; Pred. No. 1.6e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TYPNKGTOOYTDOIERPLM 20
DB 400 TYPNKGTOOYTDOIERPLM 419

```

```

RESULT 5
US-09-134-001C-3701
; Sequence 3701, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: CTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3701
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3701

```

```

Query Match 40.98; Score 45; DB 4; Length 383;
Best Local Similarity 40.0%; Pred. No. 9.6;
Matches 12; Conservative 1; Mismatches 5; Indels 12; Gaps 1;
QY 2 YFPNKGTOO-----YTDOIERPL 19
DB 212 YFPNKGTOO-----YTDOIERPL 241

```

```

RESULT 6
US-08-677-862-2
; Sequence 2, Application US/08677862
; Patent No. 5874230
; GENERAL INFORMATION:
; APPLICANT: SONG, Ho Yeong
; APPLICANT: ROTH, Mike
; TITLE OF INVENTION: TRAF2-Associated Kinase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, 34th floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/677,862
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brezner, David J
; REGISTRATION NUMBER: 24,774
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415 781-1989
; TELEFAX: 415 398-3249
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 763 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-677-862-2

```

```

Query Match 39.58; Score 43.5; DB 2; Length 763;
Best Local Similarity 47.6%; Pred. No. 39;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;
QY 4 PNKGTOO-----YTDOIERPL 19
DB 46 PNISDQVSAISYSDQIOPL 66

```

```

RESULT 7
US-09-252-571-2
; Sequence 2, Application US/09252571
; Patent No. 5981250
; GENERAL INFORMATION:
; APPLICANT: SONG, Ho Yeong
; APPLICANT: ROTH, Mike
; TITLE OF INVENTION: TRAF2-Associated Kinase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, 34th floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/252,571

```

```

: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/677,862
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Brezner, David J
: REGISTRATION NUMBER: 24,774
: REFERENCE/DOCKET NUMBER: T96-005/A63613
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415 781-1989
: TELEFAX: 415 398-3249
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 763 amino acids
: TYPE: amino acid
: STRANDEDNESS: not relevant
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
:
US-09-252-571-2

```

```

1      RESULT 8
2      US-09-434-065-2
3      ; Sequence 2, Application US/09434065
4      ; Patent No. 6107074
5      ;
6      ; GENERAL INFORMATION:
7      ;
8      ; APPLICANT: SONG, Ho Yeong
9      ; APPLICANT: ROTHE, Mike
10     ; TITLE OF INVENTION: TRAV2-Associated Kinase
11     ; NUMBER OF SEQUENCES: 2
12     ;
13     ; CORRESPONDENCE ADDRESS:
14     ; ADDRESSEE: Fleh, Hohbach, Test, Albritton & Herbertt
15     ; STREET: 4 Embarcadero Center, 34th floor
16     ; CITY: San Francisco
17     ; STATE: CA
18     ; COUNTRY: USA
19     ; ZIP: 94111
20     ;
21     ; COMPUTER READABLE FORM:
22     ; MEDIUM TYPE: Floppy disk
23     ; COMPUTER: IBM PC compatible
24     ; OPERATING SYSTEM: PC-DOS/MS-DOS
25     ; SOFTWARE: PatentIn Release #1.0, Version #1.30
26     ;
27     ; CURRENT APPLICATION DATA:
28     ; APPLICATION NUMBER: US/09/434,065
29     ; FILING DATE:
30     ; CLASSIFICATION:
31     ; PRIOR APPLICATION DATA:
32     ; APPLICATION NUMBER: 08/677,862
33     ; FILING DATE:
34     ; ATTORNEY/AGENT INFORMATION:
35     ; NAME: Brezner, David J
36     ; REGISTRATION NUMBER: 24,774
37     ; REFERENCE/DOCKET NUMBER: T96-005/A63613
38     ; TELECOMMUNICATION INFORMATION:
39     ; TELEPHONE: 415 781-1989
40     ; TELEFAX: 415 398-3249
41     ;
42     ; INFORMATION FOR SEQ ID NO: 2:
43     ; SEQUENCE CHARACTERISTICS:
44     ; LENGTH: 763 amino acids
45     ; TYPE: amino acid
46     ; STRANDEDNESS: not relevant
47     ; TOPOLOGY: linear
48     ;
49     ; MOLECULE TYPE: peptide
50     ;
51     ; US-09-434-065-2

```

Query Match	39.5%	Score 43.5;	DB 3;	length 763;
Best Local Similarity	47.6%	Pred. No. 39;		
Matches 10;	Conservative 3;	Mismatches 3;	Indels 5;	Gaps 1
QY	4	PNKGTQQ-----	YTOPIERPL	19
Db	46	PNISDQVSALSTSDIOQPL		66

```

RESULT 9
US-08-789-275-4
: Sequence 4, Application US/08789275A
: Patent No. 6251664
: GENERAL INFORMATION:
: APPLICANT: Palleja, Estivill Xavier
: APPLICANT: Pritchard, Melaine
: APPLICANT: Vilario, Jordi Guimera
: TITLE OF INVENTION: Human Gene Sequence of the Down Syndrome Critical
: Patent No. 6251664
: TITLE OF INVENTION: Region of Human Chromosome 21, Coding for A Serine
: TITLE OF INVENTION: Threonine Protein Kinase (NMB). Expressed in the
: TITLE OF INVENTION: Neuronal Regions Affected in Down Syndrome
: FILE REFERENCE: U 01114-4
: CURRENT APPLICATION NUMBER: US/08/789,275A
: CURRENT FILING DATE: 1997-01-28
: NUMBER OF SEQ ID NOS: 6
: SEQ ID NO 4
: LENGTH: 763
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
US-08-789-275-4

```

```

Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1
QY 4 PKNGTQO-----YTDQIERPL 19
      || || |::|::|
Db 46 PNISDOVSALSDDIOPL 66

RESULT 10
US-08-789-275-5
; Sequence 5, Application US/08789275A
; Patent No. 6251664
; GENERAL INFORMATION:
; APPLICANT: Palleja, Estivill Xavier
; APPLICANT: Pritchard, Melanie
; APPLICANT: Vilaro, Jordi Guimera
; TITLE OF INVENTION: Human Gene Sequence of the Down Syndrome Critical
; Patent No. 6251664
; TITLE OF INVENTION: Region of Human Chromosome 21, Coding for A Serine-
; TITLE OF INVENTION: Phreoline Protein Kinase (MNF), Expressed in the
; TITLE OF INVENTION: Neuronal Regions Affected in Down Syndrome
; FILE REFERENCE: U 01114-4
; CURRENT APPLICATION NUMBER: US/08/789, 275A
; CURRENT FILING DATE: 1997-01-28
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 5
; LENGTH: 763
; TYPE: PRT
; ORGANISM: Rat norvegicus
; FEATURE:
;
US-08-789-275-5

Query Match 39.5%; Score 43.5; DB 4; Length 763;
Best Local Similarity 47.6%; Pred. No. 39;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1
QY 4 PKNGTQO-----YTDQIERPL 19
      || || |::|::|

```

Db 46 PNISDOQVSAISYSDQIQOPL 66

RESULT 11
US-07-731-157A-2

Sequence 2, Application US/07731157A
Patent No. 5457032

GENERAL INFORMATION:

APPLICANT: Quax, Wilhelmus J.

APPLICANT: Van der Laan, Jan M.

APPLICANT: Lenting, Herman B.M.

TITLE OF INVENTION: Mutated beta-lactam acylase genes

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:

ADDRESSEE: COOLEY GODDARD CASTRO HUDDLESON & TATUM

STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/731,157A

FILING DATE: 19910509

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: EP 90200962

FILING DATE: 18-APR-1990

ATTORNEY/AGENT INFORMATION:

NAME: RAE-VENTER PH.D., BARBARA

REGISTRATION NUMBER: 32,750

REFERENCE/DOCKET NUMBER: GBRO-027/000US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-494-7622

TELEFAX: 415-857-0663

TELEX: 380816 COOLEY PA

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 720 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-731-157A-2

Query Match

Best Local Similarity 39.1%; Score 43; DB 1; Length 720;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 680 GTHYSDQIER 690

OY 7 GTQOYTDQIER 17

11 1:111111

US-08-541-780-2

Sequence 2, Application US/08541780

Patent No. 5935831

GENERAL INFORMATION:

APPLICANT: Quax, Wilhelmus J.

APPLICANT: Van der Laan, Jan M.

APPLICANT: Lenting, Herman B.M.

TITLE OF INVENTION: Mutated beta-lactam acylase genes

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:

ADDRESSEE: COOLEY GODDARD CASTRO HUDDLESON & TATUM

STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/541,780

FILING DATE: 19910509

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/731,157

FILING DATE: 18-APR-1990

ATTORNEY/AGENT INFORMATION:

NAME: RAE-VENTER PH.D., BARBARA

REGISTRATION NUMBER: 32,750

REFERENCE/DOCKET NUMBER: GBRO-027/000US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-494-7622

TELEFAX: 415-857-0663

TELEX: 380816 COOLEY PA

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 720 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-541-780-2

Query Match

Best Local Similarity 39.1%; Score 43; DB 2; Length 720;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 680 GTHYSDQIER 690

OY 7 GTQOYTDQIER 17

11 1:111111

US-08-447-464-3

Sequence 3, Application US/08447464

Patent No. 5840842

GENERAL INFORMATION:

APPLICANT: Schlessinger, Joseph

APPLICANT: Yan, Hai

TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PROTEIN

TITLE OF INVENTION: PHOSPHOTYROSINE PHOSPHATASE-SIGMA

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/447,464

FILING DATE: 24-MAY-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/130,570

FILING DATE: 01-OCT-1993

ATTORNEY/AGENT INFORMATION:

NAME: Mistock, S. Leslie

REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7683-043
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1501 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-447-464-3

Query Match 38.2%; Score 42; DB 2; Length 1501;
Best Local Similarity 60.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YFPNKGTOQY 11
|:|:|:|:|:
Db 1057 YWPNRGTEY 1066

RESULT 14
US-08-716-679-3
Sequence 3, Application US/08716679
Patent No. 5846800
GENERAL INFORMATION:
APPLICANT: Schlessinger, Joseph
APPLICANT: Yan, Hai
TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PROTEIN
TITLE OF INVENTION: PHOSPHOTYROSINE PHOSPHATASE-SIGMA
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/716,679
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/130,570
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7683-043
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1501 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-716-679-3

Query Match 38.2%; Score 42; DB 2; Length 1501;
Best Local Similarity 60.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 2 YFPNKGTOQY 11

Db 1057 YWPNRGTEY 1066
|:|:|:|:|:

RESULT 15
US-08-348-006B-5
Sequence 5, Application US/08348006B
Patent No. 5658756
GENERAL INFORMATION:
APPLICANT: RODAN, GIDEON A.
APPLICANT: SCHMIDT, AZRIEL
APPLICANT: RUTLEDGE, SU JANE
TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
TITLE OF INVENTION: TYROSINE PHOSPHATASE
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: J. MARK HAND
STREET: 126 E. LINCOLN AVE., P.O. BOX 2000
CITY: RAHWAY
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/348,006B
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/122,032
FILING DATE: 14-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: HAND, J., MARK
REGISTRATION NUMBER: 36,545
REFERENCE/DOCKET NUMBER: 189921A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-594-3905
TELEFAX: 908-594-4720
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1911 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-348-006B-5

Query Match 38.2%; Score 42; DB 1; Length 1911;
Best Local Similarity 60.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 2 YFPNKGTOQY 11
|:|:|:|:|:
Db 1467 YWPNRGTEY 1476

Search completed: December 11, 2002, 15:27:46
Job time : 15 secs

THIS PAGE BLANK (USPTO)


```

XX CC VP2 (AAR13405) and its fragments PANSE and PAPST (AAR13407) are useful
CC as immunoassay reagents for detection of anti-B19 antibodies, e.g.
CC for diagnosis of B19 infections, determining the immune status of
CC pregnant woman, testing stored blood, or selecting positive donors
CC for prodn. of B19 hyperimmune globulin preps.
CC The VP2 fragments are expressed by plasmid pUC12VP1, which contains
CC DNA sequences generated from viral DNA by PCR using the primers
CC represented in AAQ13159-63 in pairs.
CC See also AAR13400-07 and AAR13414.
XX
SQ Sequence 370 AA:
Query Match 100.0%; Score 110; DB 12; Length 370;
Best Local Similarity 100.0%; Pred. No. 5.5e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TYFPNKGTOQYTDQIERPLM 20
227 TYFPNKGTOQYTDQIERPLM 246

RESULT 2
AAR13405
ID AAR13405 standard; Protein; 543 AA.
XX
AC AAR13405;
XX
DT 24-OCT-1991 (first entry)
XX
DE Parvo virus B19 VP2.
XX
XX Primer: PCR: globulin; PANSE; PAPST.
KW
XX
OS Synthetic.
XX
XX DE4003826-A.
XX
XX 14-AUG-1991.
XX
XX 08-FEB-1990; 90DE-4003826.
XX
XX 08-FEB-1990; 90DE-4003826.
XX
XX (MIKR-) MIKROGEN MOLEKULARB.
XX
XX Soutschek E, Moltz M;
XX
XX WPI: 1991-246423/34.
XX
PT Immunologically active parvo virus B19 peptide(s) - comprising
PT capsid protein VP1 or VP2 fragments, useful for antibody
PT detection or vaccination
XX
XX Disclosure: Fig 2-6; 22pp: German.
XX
XX VP2 and its fragments PANSE (AAR13405) and PAPST (AAR13407) are useful
XX as immunoassay reagents for detection of anti-B19 antibodies, e.g.
XX for diagnosis of B19 infections, determining the immune status of
XX pregnant woman, testing stored blood, or selecting positive donors
XX for prodn. of B19 hyperimmune globulin preps.
XX The VP2 fragments are expressed by plasmid pUC12VP1, which contains
XX DNA sequences generated from viral DNA by PCR using the primers
XX represented in AAQ13159-63 in pairs.
XX See also AAR13400-07 and AAR13414.
XX
SQ Sequence 543 AA:
Query Match 100.0%; Score 110; DB 12; Length 543;
Best Local Similarity 100.0%; Pred. No. 8.6e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TYFPNKGTOQYTDQIERPLM 20

```

```

DB 400 TYFPNKGTOQYTDQIERPLM 419
|||||
RESULT 3
AAW08987
ID AAW08987 standard; Protein; 554 AA.
XX
AC AAW08987;
XX
DT 27-FEB-1997 (first entry)
XX
DE Human parvovirus VP-2 protein.
XX
XX Human: parvovirus genome: structural gene: VP-1; VP2: arthritis;
KW non-structural protein; NS: diagnosis; vaccine; parvoviral disease;
KW erythblastemia; abortion; universal fetal hydrops; liver disease;
KW haemorrhagic fever; rheumatism; detection; IgG antibody.
XX
OS Human parvovirus.
XX
XX JP07147986-A.
XX
XX 13-JUN-1995.
XX
XX 24-SEP-1992; 92JP-0281017.
XX
XX 24-SEP-1992; 92JP-0281017.
XX
XX (DENK-) DENKA SEIKEN KK.
XX
XX (ELED) DENKI KAGAKU KOGYO KK.
XX
XX WPI: 1995-242756/32.
XX
XX N-PSDB: AAT49535.
XX
XX Human parvovirus gene coding for a polypeptide - useful for
PT developing vaccines against parvoviral diseases such as
PT erythroblastemia, haemorrhagic fever, etc.
XX
XX Claim 3; Page 7-9; 38pp: English.
XX
XX The sequences given in AAW08986 represent the parvovirus structural
XX proteins, VP-1 and VP2, and the non-structural protein, NS. The
XX genomic fragment encoding these proteins may be used for the
XX diagnosis and development of vaccines for parvoviral diseases including
XX CC erythblastemia, abortion, universal fetal hydrops, liver diseases,
XX CC haemorrhagic fever, arthritis and rheumatism. The VP-1 and VP-2
XX proteins may be used to detect parvovirus IgG antibodies.
XX
SQ Sequence 554 AA:
Query Match 100.0%; Score 110; DB 16; Length 554;
Best Local Similarity 100.0%; Pred. No. 8.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TYFPNKGTOQYTDQIERPLM 20
391 TYFPNKGTOQYTDQIERPLM 410

RESULT 4
AAW23230
ID AAW23230 standard; Protein; 554 AA.
XX
AC AAW23230;
XX
XX 26-AUG-1999 (first entry)
XX
XX Erythrovirus V9 VP2 protein.
DE Erythrovirus V9 VP2 protein.
XX
XX Erythrovirus V9; differential diagnosis; parvovirus; infection;
KW erythrovirus screening; typing; immunoassay; VP2 protein.
XX

```

```

OS Erythrovirus.
XX
XX FR2771751-A1.
XX
XX 04-JUN-1999.
XX
XX 03-DEC-1997; 97FR-0015197.
XX
XX 03-DEC-1997; 97FR-0015197.
XX
XX (ASST-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.
XX
XX Auguste V, Garbarg CA, Nguyen QT;
XX
XX WPI: 1999-349543/30.
XX
XX N-PSDB: AAX81586.
XX
XX Erythrovirus V9 and its nucleic acid sequences - can be used in the
XX diagnosis of its infections
XX
XX Claim 19; Page 57-58; 80pp; French.
XX
XX The present sequence represents an erythrovirus V9 protein.
XX Probes and primers derived from erythrovirus V9 polynucleotide
XX sequences (AAX81580) can be used for differential diagnosis of
XX erythrovirus (parvovirus) infections by a combination of
XX amplification and hybridisation assay. The probes can also be
XX used to assess susceptibility to erythrovirus infection and
XX for erythrovirus screening and typing. The antibodies can be
XX used in immunoassays for diagnosis of erythrovirus V9 infections.
XX
XX Sequence 554 AA;
XX
XX Query Match 100.0%; Score 110; DB 20; Length 554;
XX Best Local Similarity 100.0%; Pred. No. 8.8e-10;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 TYFPNKGTOOYTDOIERPLM 20
XX ||||||||||||||||
XX Db 391 TYFPNKGTOOYTDOIERPLM 410
XX
XX RESULT 5
XX AAY71231
XX ID AAY71231 standard; Protein: 756 AA.
XX
XX AAY71231;
XX
XX 08-SEP-2000 (first entry)
XX
XX Capsid protein encoded by AAV2/human parvovirus B19 chimeric vector.
XX
XX Adeno-associated virus; AAV2; human parvovirus B19; chimeric;
XX recombinant parvoviral vector; cellular tropism; cap protein;
XX capsid; gene delivery; gene therapy; VP1; VP2; VP3.
XX
XX Chimeric - Adeno associated virus serotype 2.
XX Chimeric - Human parvovirus B19.
XX
XX WO200028004-A1.
XX
XX 18-MAY-2000.
XX
XX 10-NOV-1999; 99WO-US26505.
XX
XX 10-NOV-1998; 98US-0107840.
XX
XX 10-MAR-1999; 99US-0123651.
XX
XX (UYNC-) UNIV NORTH CAROLINA.
XX
XX Rabinowitz JE, Samulski RJ, Xiao W;
XX
XX WPI: 2000-376523/32.

```

```

DR N-PSDB: AAD00833.
XX
XX Recombinant parvoviral vectors with altered packaging, tropisms and
XX immunogenic properties, useful in gene therapy protocols -
XX
XX Example 21; Page 142; 153pp; English.
XX
XX The patent discloses modified parvovirus vectors with advantageous
XX antigenic properties, packaging capabilities and cellular tropisms.
XX These vectors can be used in standard recombinant DNA protocols e.g. gene
XX therapy for delivering nucleic acids to cells.
XX
XX The present sequence is a capsid protein encoded by an
XX adeno-associated virus serotype 2 (AAV2)/human parvovirus B19 chimeric
XX vector. This vector encodes AAV2 VP1 and VP2 capsid proteins
XX and human parvovirus B19 VP2 protein. The chimeric vector was
XX constructed by replacing the VP3 major cap protein of AAV2 with B19's
XX VP2. Recombinant parvovirus comprising the chimeric capsid is useful
XX for gene delivery.
XX
XX Sequence 756 AA;
XX
XX Query Match 100.0%; Score 110; DB 21; Length 756;
XX Best Local Similarity 100.0%; Pred. No. 1.3e-09;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 TYFPNKGTOOYTDOIERPLM 20
XX ||||||||||||||||
XX Db 593 TYFPNKGTOOYTDOIERPLM 612
XX
XX RESULT 6
XX AAM08986
XX ID AAM08986 standard; Protein: 781 AA.
XX
XX AAM08986;
XX
XX 27-FEB-1997 (first entry)
XX
XX Human parvovirus VP-1 protein.
XX
XX Human; parvovirus genome; structural gene; VP-1; VP2; arthritis;
XX non-structural protein; NS; diagnosis; vaccine; parvoviral disease;
XX erythblastemia; abortion; universal fetal hydrops; liver disease;
XX haemorrhagic fever; rheumatism; detection; IgG antibody.
XX
XX Human parvovirus.
XX
XX Key Location/Qualifiers
XX Misc-difference 19
XX /label= Gly, Val
XX Misc-difference 61
XX /label= Asn, Asp
XX Misc-difference 220
XX /label= His, Asn
XX
XX JP07147986-A.
XX
XX 13-JUN-1995.
XX
XX 24-SEP-1992; 92JP-0281017.
XX
XX 24-SEP-1992; 92JP-0281017.
XX
XX (DENK-) DENKA SEIKEN KK.
XX (ELED) DENKI KAGAKU KOGYO KK.
XX
XX WPI: 1995-242756/32.
XX
XX N-PSDB: AAT49535.
XX
XX Human parvovirus gene coding for a polypeptide - useful for
XX developing vaccines against parvoviral diseases such as
XX erythroblastemia, haemorrhagic fever, etc.

```

PS Claim 2; Page 5-7; 38pp; English.

XX The sequences given in AAM08986 represent the parvovirus structural

CC proteins, VP-1 and VP2, and the non-structural protein, NS. The

CC genomic fragment encoding these proteins may be used for the

CC diagnosis and development of vaccines for parvoviral diseases including

CC erythblastemia, abortion, universal fetal hydrops, liver diseases,

CC hemorrhagic fever, arthritis and rheumatism. The VP-1 and VP-2

CC proteins may be used to detect parvovirus IgG antibodies.

XX

SO Sequence 781 AA;

Query Match 100.0%; Score 110; DB 16; Length 781;

Best Local Similarity 100.0%; Pred. No. 1.3e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYPNKGTOQYTDQIERPLM 20

DB 618 TYPNKGTOQYTDQIERPLM 637

RESULT 7

ID AAY23227 standard; Protein: 781 AA.

XX AAY23227;

AC AAY23227;

XX 26-AUG-1999 (first entry)

DT Erythrovirus V9 VP1 protein.

XX

DE Erythrovirus V9 VP1 protein.

XX

XX Erythrovirus V9; differential diagnosis; parvovirus; infection;

KM erythrovirus screening; typing; immunoassay; VP1 protein.

XX

OS Erythrovirus.

XX

XX FR2771751-A1.

PN

PD 04-JUN-1999.

XX

PF 03-DEC-1997; 97FR-0015197.

XX

PR 03-DEC-1997; 97FR-0015197.

XX

PA (ASSI-) ASSISTANCE PUBLIQUE HOPITALUX PARIS.

XX

PI Auguste V, Garbarg CA, Nguyen QT;

XX

XX WPI; 1999-349543/30.

XX

XX N-PSDB; AAX81583.

XX

PT Erythrovirus V9 and its nucleic acid sequences - can be used in the

XX diagnosis of its infections

XX

PS Claim 19; Page 50-52; 80pp; French.

XX

XX The present sequence represents an erythrovirus V9 protein.

CC Probes and primers derived from erythrovirus V9 polynucleotide

CC sequences (AAX81580) can be used for differential diagnosis of

CC erythrovirus (parvovirus) infections by a combination of

CC amplification and hybridisation assay. The probes can also be

CC used to assess susceptibility to erythrovirus infection and

CC for erythrovirus screening and typing. The antibodies can be

CC used in immunoassays for diagnosis of erythrovirus V9 infections.

XX

SO Sequence 781 AA;

Query Match 100.0%; Score 110; DB 20; Length 781;

Best Local Similarity 100.0%; Pred. No. 1.3e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYPNKGTOQYTDQIERPLM 20

DB 618 TYPNKGTOQYTDQIERPLM 637

DB 618 TYPNKGTOQYTDQIERPLM 637

RESULT 8

ID AAG83008 standard; Protein: 383 AA.

XX AAG83008;

AC AAG83008;

XX

DT 03-SEP-2001 (first entry)

XX

DE S. epidermidis open reading frame protein sequence SEQ ID NO:3110.

XX

XX Staphylococcus epidermidis SRI strain; infection; diagnosis;

KM vaccination; endocarditis.

XX

OS Staphylococcus epidermidis.

XX

XX WO200134809-A2.

PN

PD 17-MAY-2001.

XX

PF 09-NOV-2000; 2000WO-US30782.

XX

PR 09-NOV-1999; 99US-0164258.

XX

PA (GLAX) GLAXO GROUP LTD.

XX

PI Kimmery MJ;

XX

XX WPI; 2001-316495/33.

DR N-PSDB; AAH53858.

XX

PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,

XX useful for vaccinating against infections, e.g. endocarditis -

XX

PS Claim 18; Page 820; 2188pp; English.

XX

XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides

CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.

CC (I) and (II) can have antibacterial activity and therefore can be used

CC in vaccination. The nucleic acids (I) may be used to produce the

CC S. epidermidis polypeptides (II) via the production of vectors

CC containing them which are used to produce hosts cells which express the

CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be

CC used to vaccinate subjects and to raise antibodies against the bacteria.

CC The polypeptides may also be used to assay for other inhibitors of their

CC activity and therefore identify compounds that may be used for the

CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to

CC AAH55090 represent specifically claimed S. epidermidis genomic DNA

CC polynucleotide sequences from the present invention. AAH55091 to

CC AAH55098 represent oligonucleotide sequences and primers which are used

CC in the exemplification of the present invention.

CC N.B. The present invention specifically claims all the polynucleotide

CC sequences given in the sequence listing of the present specification,

CC however the sequence listing only goes up to SEQ ID NO:4454 so even

CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,

CC no sequences are present for SEQ ID NO:4455 to 4464.

XX

SO Sequence 383 AA;

Query Match 40.9%; Score 45; DB 22; Length 383;

Best Local Similarity 40.0%; Pred. No. 47;

Matches 12; Conservative 1; Mismatches 5; Indels 12; Gaps 1;

OY 2 YFPNKGTOQ-----YTDQIERPL 19

DB 212 YGPQGAQOKMIPKLDLSALRHYHDKIEREL 241

RESULT 9

ABP38856

ID ABP38856 standard; Protein: 383 AA.

XX ABP38856;
XX
XX 24-JUL-2002 (first entry)
XX
DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3701.
XX
XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
XX antibacterial; gene therapy.
XX
XX Staphylococcus epidermidis.
XX
XX US6380370-B1.
XX
XX 30-APR-2002.
XX
XX 13-AUG-1998; 98US-0134001.
XX
XX 14-AUG-1997; 97US-055779P.
XX 08-NOV-1997; 97US-064964P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Doucette-Stamm LA, Bush D;
XX
XX WPI: 2002-381255/41.
XX N-PSDB; ABN91401.
XX
XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis
XX polypeptide, useful for diagnosing and treating bacterial infections -
XX
XX Disclosure; SEQ ID 3701; 267pp; English.
XX
XX
XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
XX frame (ORF) nucleic acid sequences which encode the amino acid sequences
XX given in ABP5124 to ABP37960. The S. epidermidis sequences have
XX antibacterial activity and can be used in gene therapy. The sequences
XX can also be used in the diagnosis and treatment of bacterial infections,
XX particularly S. epidermidis infections. The sequences can be used to
XX screen for compounds able to interfere with the S. epidermidis life
XX cycle or inhibit S. epidermidis infection.
XX N.B. The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from the
XX USPTO web site.
XX
XX
SQ Sequence 383 AA;
Query Match 40.9%; Score 45; DB 23; Length 383;
Best Local Similarity 40.0%; Pred. No. 47;
Matches 12; Conservative 1; Mismatches 5; Indels 12; Gaps 1;
OY 2 YFPKGTQO-----YTDQIERPL 19
DB 212 YGPKGAQOKMIPKIDSALRHYHDKIEREL 241
RESULT 10
ABG19946
ID ABG19946 standard; Protein: 767 AA.
XX
XX ABG19946;
XX
XX 18-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #19937.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX
XX

PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HNSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI: 2001-639362/73.
XX N-PSDB; AAS84133.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX
XX Claim 20; SEQ ID No 50305; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human
XX diagnostic amino acid sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 767 AA;
Query Match 40.9%; Score 45; DB 22; Length 767;
Best Local Similarity 53.8%; Pred. No. 1,1e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
OY 3 FPNKGTQOYTDQI 15
DB 511 FPNKGTSTYADQM 523
RESULT 11
ABG13993
ID ABG13993 standard; Protein: 904 AA.
XX
XX ABG13993;
XX
XX 18-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #13984.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX
XX

PF 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Dmanac RT, Liu C, Tang YT;
PI
XX
XX WPI; 2001-639362/73.
DR
DR N-PSDB; AAS78180.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
PS Claim 20; SEQ ID No 44352; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pcl_sequences](http://wipo.int/pub/published_pcl_sequences).

```

Query Match      40.0%:  Score 44:  DB 22:  Length 904;
Best Local Similarity 46.7%:  Pred. No. 1.9e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0.

```

XX	RESULT 12
XX	ABB28188
ID	ABB28188 standard; Peptide; 50 AA.
AC	ABB28188;
XX	
DT	01-FEB-2002 (first entry)
DE	
XX	Human peptide #839 encoded by breast cell single exon nucleic acid probe
XX	
KW	Human; microarray; single exon probe; gene expression; breast;
KW	disease; cancer.
XX	
OS	Homo sapiens.
XX	
PN	WO200157271-A2.
XX	
PD	09-AUG-2001.
XX	
PF	30-JAN-2001; 2001WO-US00662.
XX	

PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0633365.
PR 12-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC

PI	Penn SG, Hanzel DK, Chen W, Rank DR,
XX	
DR	WPI; 2001-496933/54.

XX PT New spatially-addressable set of single exon nucleic acid probes,
PS useful for measuring gene expression in sample derived from human
XX PT breast, comprises number of single exon nucleic acid probes -
Claim 27; SEQ ID NO 11156; 327pp + sequence listing; English.

CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human breast and BT 474 cells. The method involves contacting
 CC the probes with a collection of detectably labelled nucleic acids
 CC derived from mRNA of human breast, and then measuring the label
 CC bound to each probe of the microarray. The probes are useful for
 CC verifying the expression of regions of genomic DNA predicted to
 CC encode proteins. They are useful for gene discovery, and for
 CC determining predisposition and/or prognosing breast disease. Gene
 CC expression analysis is useful for assessing the toxicity of chemical
 CC agents on cells. The microarray of this invention presents a far greater
 CC diversity of probes for measuring gene expression, with far less bias
 CC than expressed sequence tag microarrays. The method is suitable for
 CC rapid production of functional information from genomic sequence. The
 CC present sequence is a peptide encoded by a single exon nucleic acid
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 CC Sequence 50 AA:
 CC
 CC

Query Match	39.5%	Score 43.5	DB 22	Length 50
Best Local Similarity	47.6%	Pred. No. 7.4		
Matches 10; Conservative	3;	Mismatches	3;	Indels 5; Gaps 1
QY	4	PNKGTQ-----YTDIERPL	19	
Db	27	PNISDQVSAISYSIDQIQPL	47	

	RESULT 13
ABB33363	
ID	ABB33363 standard; Peptide: 50 AA.
XX	
AC	ABB33363;
XX	
DT	04-FEB-2002 (first entry)
XX	
DE	Peptide #669 encoded by human foetal liver single exon probe.
XX	
KE	Human; foetal liver; gene expression; single exon nucleic acid probe.
XX	
OS	Homo sapiens.
XX	
PN	WO200157277-A2.
XX	
PD	09-AUG-2001.
XX	
PF	30-JAN-2001; 2001WO-US00669.
XX	
PR	04-FEB-2000; 2000US-0180312.
RR	26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
PS Claim 27; SEQ ID NO 25998; 639bp + sequence listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pcr_sequences.
XX
SQ Sequence 50 AA;

Query Match 39.5%; Score 43.5; DB 22; Length 50;
Best Local Similarity 47.6%; Pred. No. 7.4;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

OY 4 PNKGTOQ-----YTQIERPL 19
II II I:||||:II
DB 27 PNISDQVSALSYPDIOQPL 47

RESULT 14

ABBI8822
ID ABBI8822 standard; Protein; 50 AA.

XX
AC ABBI8822;

XX
DT 23-JAN-2002 (first entry)

XX
DE Protein #821 encoded by probe for measuring heart cell gene expression.

XX
KW Human: gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.

XX
OS Homo sapiens.

XX
PN WO200157274-A2.

XX
PD 09-AUG-2001.

XX
PE 30-JAN-2001; 2001WO-US000666.

XX
PR 04-FEB-2000; 2000US-0180312.

XX
PR 26-MAY-2000; 2000US-0207456.

XX
PR 30-JUN-2000; 2000US-0608408.

XX
PR 03-AUG-2000; 2000US-0632366.

XX
PR 21-SEP-2000; 2000US-0234687.

XX
PR 27-SEP-2000; 2000US-0236359.

XX
PR 04-OCT-2000; 2000GB-0024263.

XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX
DR WPI; 2001-488899/53.

XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
XX Claim 15; SEQ ID NO 20592; 530bp; English.

XX
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcr_sequences.

XX
SQ Sequence 50 AA;

Query Match 39.5%; Score 43.5; DB 22; Length 50;
Best Local Similarity 47.6%; Pred. No. 7.4;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

OY 4 PNKGTOQ-----YTQIERPL 19
II II I:||||:II
DB 27 PNISDQVSALSYPDIOQPL 47

RESULT 15

AAM54148
ID AAM54148 standard; Protein; 50 AA.

XX
AC AAM54148;

XX
DT 05-NOV-2001 (first entry)

XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 26253.

XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.

XX
OS Homo sapiens.

XX
PN WO200157275-A2.

XX
PD 09-AUG-2001.

XX
PE 30-JAN-2001; 2001WO-US000667.

XX
PR 04-FEB-2000; 2000US-0180312.

XX
PR 26-MAY-2000; 2000US-0207456.

XX
PR 30-JUN-2000; 2000US-0608408.

XX
PR 03-AUG-2000; 2000US-0632366.

XX
PR 21-SEP-2000; 2000US-0234687.

XX
PR 27-SEP-2000; 2000US-0236359.

XX
PR 04-OCT-2000; 2000GB-0024263.

XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX
DR WPI; 2001-483446/52.

XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -

XX
PS Example 4; SEQ ID NO: 26253; 650bp + Sequence Listing; English.

XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.

XX
 SO Sequence 50 AA;

Query Match 39.5%; Score 43.5; DB 22; Length 50;
 Best Local Similarity 47.6%; Pred. No. 7.4;
 Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

OY 4 PNKGIGQ-----YTDQIERPL 19
 |||||
 |||||
 Db 27 PNISDQVSAISLSPDIQQL 47

Search completed: December 11, 2002, 15:26:10
 Job time : 37 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 15:24:57 : Search time 29 Seconds
(without alignments)
142.101 Million cell updates/sec

Title: US-09-991-433-48
Perfect score: 110
Sequence: 1 TYPFNKGTQGYTDIERPLM 20

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriaph:*
17: SP archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	110	100.0	138	12	085155 human parvo
2	110	100.0	141	12	085146 human parvo
3	110	100.0	141	12	085166 human parvo
4	110	100.0	141	12	085171 human parvo
5	110	100.0	142	12	085168 human parvo
6	110	100.0	142	12	085173 human parvo
7	110	100.0	144	12	085138 human parvo
8	110	100.0	145	12	085161 human parvo
9	110	100.0	145	12	085181 human parvo
10	110	100.0	146	12	085158 human parvo
11	110	100.0	147	12	085142 human parvo
12	110	100.0	148	12	085150 human parvo
13	110	100.0	151	12	085131 human parvo
14	110	100.0	151	12	085177 human parvo
15	110	100.0	153	12	085123 human parvo
16	110	100.0	157	12	085196 human parvo

17	110	100.0	162	12	085135 human parvo
18	110	100.0	546	12	0913X0 human parvo
19	110	100.0	546	12	0913W7 human parvo
20	110	100.0	554	12	09JCEP7 human parvo
21	110	100.0	554	12	09WKL9 human parvo
22	110	100.0	554	12	09PZS9 human parvo
23	110	100.0	554	12	0912B7 human parvo
24	110	100.0	554	12	065790 human parvo
25	110	100.0	554	12	090201 human parvo
26	110	100.0	760	12	09PZT8 human parvo
27	110	100.0	761	12	09PZU0 human parvo
28	110	100.0	765	12	09PZT6 human parvo
29	110	100.0	769	12	09PZT4 human parvo
30	110	100.0	773	12	0913X1 human parvo
31	110	100.0	773	12	0913W8 human parvo
32	110	100.0	781	12	09JES0 human parvo
33	110	100.0	781	12	09JGP8 human parvo
34	110	100.0	781	12	09WKM0 human parvo
35	110	100.0	781	12	09PZT0 human parvo
36	110	100.0	781	12	0912B8 human parvo
37	110	100.0	781	12	089316 human parvo
38	110	100.0	781	12	089317 human parvo
39	110	100.0	781	12	089318 human parvo
40	110	100.0	781	12	089319 human parvo
41	110	100.0	781	12	089320 human parvo
42	110	100.0	781	12	089321 human parvo
43	110	100.0	781	12	090221 human parvo
44	110	100.0	781	12	090222 human parvo
45	110	100.0	781	12	090223 human parvo

ALIGNMENTS

```

RESULT 1
ID 085155 PRELIMINARY; PRT; 138 AA.
AC 085155;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemanuer A., Von Pohlitzki A., Gigler A., Cassinotti P., Siegl G.,
  Wolf H., Modrow S.;
  "XXXSequence variability among different parvovirus B19 isolates.";
  Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RL EMBL; Z70564; CA94477.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15237 MW; 5899F8879A3E68D CRC64;
  Query Match 100.0%; Score 110; DB 12; Length 138;
  Best Local Similarity 100.0%; Pred. No. 1.9e-10;
  Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 1 TYPFNKGTQGYTDIERPLM 20
  Db 103 TYPFNKGTQGYTDIERPLM 122
  RESULT 2
  ID 085146 PRELIMINARY; PRT; 141 AA.
  AC 085146;
  DT 01-NOV-1996 (TREMblrel. 01, Created)

```


RA Wolf H., Modrow S.;
 RT "XXXsequence variability among different parvovirus B19 isolates.";
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z70582; CA94495.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON_TER 1
 FT NON_TER 142 142
 SQ SEQUENCE 142 AA; 15820 MW; 3A9E5E07C0BC0434 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 142;
 Best Local Similarity 100.0%; Pred. No. 2e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDOIERPLM 20
 Db 101 TYFPNKGTOOYTDOIERPLM 120

RESULT 7

085138
 ID 085138 PRELIMINARY; PRT; 144 AA.
 AC 085138;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauner A., Von Poblotszki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXsequence variability among different parvovirus B19 isolates.";
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z70547; CA94459.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON_TER 1
 FT NON_TER 144 144
 SQ SEQUENCE 144 AA; 16022 MW; 95D6EF37BA5997AE CRC64;

Query Match 100.0%; Score 110; DB 12; Length 144;
 Best Local Similarity 100.0%; Pred. No. 2e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDOIERPLM 20
 Db 103 TYFPNKGTOOYTDOIERPLM 122

RESULT 8

085161
 ID 085161 PRELIMINARY; PRT; 145 AA.
 AC 085161;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauner A., Von Poblotszki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXsequence variability among different parvovirus B19 isolates.";
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z70570; CA94483.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.

FT NON_TER 1
 FT NON_TER 145 145
 SQ SEQUENCE 145 AA; 16165 MW; B4FE95AB6EAD9A67 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 145;
 Best Local Similarity 100.0%; Pred. No. 2e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDOIERPLM 20
 Db 101 TYFPNKGTOOYTDOIERPLM 120

RESULT 9

085181
 ID 085181 PRELIMINARY; PRT; 145 AA.
 AC 085181;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauner A., Von Poblotszki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXsequence variability among different parvovirus B19 isolates.";
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z70590; CA94503.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON_TER 1
 FT NON_TER 145 145
 SQ SEQUENCE 145 AA; 16210 MW; 6B45A1E1B9E923C4 CRC64;

Query Match 100.0%; Score 110; DB 12; Length 145;
 Best Local Similarity 100.0%; Pred. No. 2e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDOIERPLM 20
 Db 105 TYFPNKGTOOYTDOIERPLM 124

RESULT 10

085158
 ID 085158 PRELIMINARY; PRT; 146 AA.
 AC 085158;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauner A., Von Poblotszki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXsequence variability among different parvovirus B19 isolates.";
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z70567; CA94480.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON_TER 1
 FT NON_TER 146 146
 SQ SEQUENCE 146 AA; 16252 MW; 2B74FE95AB6EAD9A CRC64;

Query Match 100.0%; Score 110; DB 12; Length 146;

Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYPFNKGTQOYTDOIERPLM 20
DB 101 TYPFNKGTQOYTDOIERPLM 120

RESULT 11

ID 085142 PRELIMINARY; PRT: 147 AA.

AC 085142; 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.

OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;

RP SEQUENCE FROM N.A.
[1]

RA Hemauer A., Von Poblotski A., Gigler A., Cassinotti P., Siegl G.,
Wolff H., Modrow S.;
RT "XXSsequence variability among different parvovirus B19 isolates."
RL EMBL: 270551; CA94463.1; -
DR InterPro: IPR001403; Parvo_coat.

DR Pfam: PF00740; Parvo_coat: 1.
FT NON_TER 1 1
FT SEQUENCE 147 AA: 16424 MW; E4ECF2459B308BA9 CRC64;

SO Query Match 100.0%; Score 110; DB 12; Length 147;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYPFNKGTQOYTDOIERPLM 20
DB 104 TYPFNKGTQOYTDOIERPLM 123

RESULT 12

ID 085150 PRELIMINARY; PRT: 148 AA.

AC 085150; 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.

OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;

RP SEQUENCE FROM N.A.
[1]

RA Hemauer A., Von Poblotski A., Gigler A., Cassinotti P., Siegl G.,
Wolff H., Modrow S.;
RT "XXSsequence variability among different parvovirus B19 isolates."
RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: 270553; CA94471.1; -

DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat: 1.
FT NON_TER 1 1
FT SEQUENCE 148 AA: 16539 MW; 881AECF2459B308B CRC64;

SO Query Match 100.0%; Score 110; DB 12; Length 148;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYPFNKGTQOYTDOIERPLM 20
DB 105 TYPFNKGTQOYTDOIERPLM 124

DB 104 TYPFNKGTQOYTDOIERPLM 123

RESULT 13

ID 085131 PRELIMINARY; PRT: 151 AA.

AC 085131; 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.

OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;

RP SEQUENCE FROM N.A.
[1]

RX MEDLINE-96332516; PubMed-8760426;
RA Hemauer A., Von Poblotski A., Gigler A., Cassinotti P., Siegl G.,
Wolff H., Modrow S.;
RT "Sequence variability among different parvovirus B19 isolates."
RL J. Gen. Virol. 77:1781-1785(1996).

DR EMBL: 270540; CA94452.1; -
DR InterPro: IPR001403; Parvo_coat.

DR Pfam: PF00740; Parvo_coat: 1.
FT NON_TER 1 1
FT SEQUENCE 151 AA: 16902 MW; 3CA74914B8E73A3E CRC64;

SO Query Match 100.0%; Score 110; DB 12; Length 151;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYPFNKGTQOYTDOIERPLM 20
DB 105 TYPFNKGTQOYTDOIERPLM 124

RESULT 14

ID 085177 PRELIMINARY; PRT: 151 AA.

AC 085177; 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.

OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;

RP SEQUENCE FROM N.A.
[1]

RA Hemauer A., Von Poblotski A., Gigler A., Cassinotti P., Siegl G.,
Wolff H., Modrow S.;
RT "XXSsequence variability among different parvovirus B19 isolates."
RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: 270586; CA94499.1; -

DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat: 1.
FT NON_TER 1 1
FT SEQUENCE 151 AA: 16888 MW; 06C6A614A616DDA3 CRC64;

SO Query Match 100.0%; Score 110; DB 12; Length 151;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYPFNKGTQOYTDOIERPLM 20
DB 105 TYPFNKGTQOYTDOIERPLM 124

RESULT 15

```

085123
ID 085123 PRELIMINARY; PRT; 153 AA.
AC 085123;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN vpi.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96332516; PubMed=8760426;
RA Hemaue A., Von Poblitzki A., Giegler A., Cassinotti P., Siegl G.,
  Wolf H., Modrow S.;
RT " Sequence variability among different parvovirus B19 isolates.";
RL J. Gen. Virol. 77:1781-1785(1996).
DR EMBL; Z70532; CAA94444.1;
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1
FT NON_TER 153
SQ SEQUENCE 153 AA; 17131 MW; ECCA1F44020814EC CRC64;

Query Match 100.0%; Score 110; DB 12; Length 153;
Best Local Similarity 100.0%; Pred. No. 2,1e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQIERPIIM 20
   |||||
Db 104 TYFPNKGTOQYTDQIERPIIM 123

```

Search completed: December 11, 2002, 15:27:02
 Job time : 30 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 15:24:27 ; Search time 10 Seconds
(without alignments)
82.953 Million cell updates/sec

Title: US-09-991-433-48
Perfect score: 110
Sequence: 1 TYPNKGTOYTQDIERPLM 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	781	1 COAT_PAVHB	P07299 human parvo
2	51	46.4	458	1 NU4M_STRCA	O21406 struthio ca
3	49	44.5	1856	1 MGA_HUMAN	O43451 homo sapien
4	48	43.6	1336	1 SEQ3_YEAST	P33332 saccharomyc
5	44	40.0	508	1 VGEG_HINV	P07923 infectious
6	44	40.0	547	1 CH60_PASMO	O59687 pasteurella
7	44	40.0	548	1 CH60_HAEIN	P43753 haemophilus
8	43.5	39.5	763	1 DYRA_HUMAN	O13627 homo sapien
9	43.5	39.5	763	1 DYRA_RAT	O63470 rattus norv
10	43	39.1	546	1 CH60_ACPPL	P94166 actinobacil
11	43	39.1	720	1 G7AC_BREDI	O91586 brevundimon
12	43	39.1	913	1 UBPO_HUMAN	O9bxu7 homo sapien
13	42.5	38.6	248	1 UBIG_RHIME	O92mk1 rhizobium m
14	42	38.2	765	1 BGLX_SALTY	O56078 salmonella
15	41	37.3	266	1 MURI_STEAM	O99uv6 staphylococ
16	41	37.3	286	1 SPEE_BUCAT	P57305 buchiera ap
17	41	37.3	306	1 PYRD_METJA	O58070 methanococ
18	40.5	36.8	349	1 DDL_LACIA	O9c115 lactococcus
19	40	36.4	302	1 CANI_RABIT	O68815 oryctolagus
20	40	36.4	459	1 NU4M_CHICK	P18939 gallus gall
21	40	36.4	713	1 CANI_MOUSE	O35350 mus musculu
22	40	36.4	713	1 CANI_RAT	P97571 rattus norv
23	40	36.4	725	1 GUNG_CIOCE	P37700 clostridium
24	40	36.4	864	1 LOXX_SOTBN	P24095 glycyne max
25	40	36.4	1066	1 HYSA_STRPN	O54873 streptococ
26	40	36.4	1257	1 PER2_MOUSE	O54943 mus musculu
27	40	36.4	1826	1 T2DI_HUMAN	P07768 oryctolagus
28	40	36.4	1872	1 T2DI_HUMAN	P21675 homo sapien
29	39.5	35.9	151	1 RS9_MYCTU	O06259 mycobacteri
30	39.5	35.5	153	1 RS9_MYCTU	P40828 mycobacteri
31	39	35.5	141	1 HBAD_CATMA	P82112 catharacta
32	39	35.5	141	1 HBAD_PASMO	P07413 passer mont
33	39	35.5	141	1 HBAD_STRUVU	F02004 sturnus vul

34	39	35.5	245	1 IE0_NPVOP	O10369 oryza pseu
35	39	35.5	326	1 VS09_ROT99	P33492 human rotav
36	39	35.5	398	1 SPEE_STRPY	P00788 streptococ
37	39	35.5	414	1 YAFB_ECOLI	P04335 escherichia
38	39	35.5	424	1 HDAD3_MOUSE	O88895 mus musculu
39	39	35.5	428	1 HDAD3_CHICK	P56520 gallus gall
40	39	35.5	428	1 HDAD3_HUMAN	O15379 homo sapien
41	39	35.5	509	1 PR12_HUMAN	P49643 homo sapien
42	39	35.5	545	1 GVPD_HALME	O02229 halobacteri
43	39	35.5	546	1 VLI_HPV19	O02050 human papil
44	39	35.5	581	1 VG25_ARCFU	O28658 archaeoglob
45	39	35.5	708	1 EFG1_STRCO	P40173 streptomyc

ALIGNMENTS

RESULT 1					
ID	COAT_PAVHB	STANDARD	PRT	781 AA.	
AC	P07299;				
DT	01-APR-1988 (Rel. 07, Created)				
DT	01-APR-1988 (Rel. 07, Last sequence update)				
DT	01-APR-1990 (Rel. 14, Last annotation update)				
DE	Probable coat protein VPI.				
OS	Human parvovirus B19.				
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.				
OX	NCBI_TaxID=10798;				
FM	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Isolate AU;				
RX	MEDLINE=86200451; Pubmed=3701931;				
RA	Shade R.O., Blundell M.C., Colmore S.F., Tattersall P., Astell C.R.;				
RT	"Nucleotide sequence and genome organization of human parvovirus B19				
RT	isolated from the serum of a child during aplastic crisis.";				
RL	J. Virol. 58:921-936(1986).				
CC	-----				
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; M13178; AAA66867.1; -				
DR	PIR; A24299; VCPV19.				
DR	InterPro; IPR001403; Parvo_coat.				
DR	Pfam; PF00740; Parvo_coat; 1.				
KW	Coat protein; Glycoprotein.				
FT	CARBOHYD 46				
FT	CARBOHYD 184				
FT	CARBOHYD 220				
FT	CARBOHYD 293				
SO	SEQUENCE 781 AA; 86015 MW; 806254DSD0516807 CRC64;				
Query Match					
Best Local Similarity 100.0%; Score 110; DB 1; Length 781;					
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
OY	1 TYPNKGTOYTQDIERPLM 20				
DB	618 TYPNKGTOYTQDIERPLM 637				
RESULT 2					
ID	NU4M_STRCA	STANDARD	PRT	458 AA.	
AC	O21406;				
DT	15-JUL-1998 (Rel. 36, Created)				
DT	15-JUL-1998 (Rel. 36, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	NMH-ubiquitinone oxidoreductase chain 4 (EC 1.6.5.3).				

```

GN MTND4 OR ND4 OR NADH4.
OS Struthio camelus (Ostrich).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Palaeognathae; Struthioniformes; Struthionidae;
OC Struthio.
OX NCBI_TaxID=9801;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97357422; PubMed=9214748;
RA Harlid A., Janke A., Arnason U.;
RT "The mtDNA sequence of the ostrich and the divergence between
RL palaeognathous and neognathous birds.";
RN Møl. Biol. Evol. 14:734-761(1997).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21263106; PubMed=11370967;
RA Haddrath O., Baker A.J.;
RT "Complete mitochondrial DNA genome sequences of extinct birds: rattle
RN phylogenetics and the vicariance biogeography hypothesis.";
RX Proc. R. Soc. Lond., B, Biol. Sci. 268:939-945(2001).
CC -I- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on ways
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, Y12025; CAA72753.1; -.
DR EMBL, AF338715; AAK53354.1; -.
DR InterPro: IPR003918; MADhub_oxred4.
DR InterPro: IPR001750; Oxidored_q1.
DR InterPro: IPR000260; Oxidored_q5_N.
DR Pfam: PF00361; Oxidored_q1_1.
DR Pfam: PF01059; Oxidored_q5_N; 1.
DR PRINTS: PR01437; NDOXDRPTASE4.
DR Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 458 AA; 50955 MW; 575118B6297A18AC CRC64;

```

RT	to sucrose-isomaltase.";	
RL	J. Biol. Chem. 273:3076-3081(1998).	
RN	(2)	
RP	REVISIONS TO 776: 1049; 1100; 1541; 1612 AND 1811.	
RA	Nichols B.L., Eldering J.A., Avery S.E., Hahn D., Quaroni A.,	
RA	Sterchi E.E.;	
RL	Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.	
RN	[3]	
RP	CHARACTERIZATION.	
RC	TISSUE-Small intestine mucosa;	
RX	Medline-89066802; PubMed-3143729;	
RT	Naim H.Y., Sterchi E.E., Lentez M.J.;	
RA	"Structure, biosynthesis, and glycosylation of human small intestinal	
RT	maltase-glucosylase."	
RL	J. Biol. Chem. 263:19709-19717(1988).	
RN	[4]	
RP	SULFATION.	
RX	Medline-88082658; PubMed-3121301;	
RA	Danielson E.M.;	
RT	"Tyrosine sulfation, a post-translational modification of microvillar	
RT	enzymes in the small intestinal enterocyte.";	
RL	EMBO J. 6:2891-2896(1987).	
CC	-1- FUNCTION: MAY SERVE AS AN ALTERNATE PATHWAY FOR STARCH DIGESTION	
CC	WHEN LUMINAL ALPHA-AMYLASE ACTIVITY IS REDUCED BECAUSE OF	
CC	IMMATUREITY OR MALNUTRITION. MAY PLAY A UNIQUE ROLE IN THE	
CC	DIGESTION OF MALTED DIETARY OLIGOSACCHARIDES USED IN FOOD	
CC	MANUFACTURING.	
CC	-1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing 1,4-	
CC	linked D-glucose residues with release of D-glucose.	
CC	-1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-	
CC	glucose residues successively from non-reducing ends of the chains	
CC	with release of beta-D-glucose.	
CC	-1- SUBUNIT: MONOMER.	
CC	-1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. BRUSH BORDER.	
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN SMALL INTESTINE, GRANULOCYTE, AND	
CC	KIDNEY BUT NOT IN SALIVARY GLAND OR PANCREAS.	
CC	-1- PTM: N- AND O-GLYCOSYLATED.	
CC	-1- PTM: DOES NOT UNDERGO INTRACELLULAR OR EXTRACELLULAR PROTEOLYTIC	
CC	CLEAVAGE.	
CC	-1- PTM: SULFATED (BY SIMILARITY).	
CC	-1- SIMILARITY: BELONGS TO FAMILY 31 OF GLYCOSYL HYDROLASES.	
CC	-1- SIMILARITY: CONTAINS 2 P-TYPE (TREFOL) DOMAINS.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; AF016833; AAC39568.2; .	
DR	Genew; HGNC:7043; MGAM.	
DR	MM: 154360. .	
DR	InterPro; IPR000322; Glyco_hydro_31.	
DR	InterPro; IPR000519; P_trefol1.	
DR	Pfam; PF000088; trefol1.2.	
DR	Pfam; PF01055; Glyco_hydro_31.2.	
DR	SMART; SMO0018; P: 2	
DR	PROSITE; PS00129; GLYCOSYL-HYDROL_F31_1; 2.	
DR	PROSITE; PS00025; P_TREFOL1; 1.	
KW	Multifunctional enzyme; Transmembrane; Glycoprotein; Hydrolase;	
KW	Glycosidase; Repeat; Signal-anchor; Sulfation.	
FT	INT MET 0	
FT	DOMAIN 1 12	CYTOPLASMIC (POTENTIAL).
FT	TRANSHEM 13 33	SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT		(POTENTIAL).
FT	DOMAIN 34 1856	LUMENAL (POTENTIAL).
FT	DOMAIN 37 83	SER/THR-RICH.
FT	DOMAIN 88 132	P-TYPE 1.
FT	DOMAIN 197 914	MALTASE.
FT	DOMAIN 952 998	P-TYPE 2.
FT	DOMAIN 1066 1812	GLUCOAMYLASE.


```

FT ACT_SITE 528 528 BY SIMILARITY.
FT ACT_SITE 1419 1419 BY SIMILARITY.
FT DISULFID 89 117 BY SIMILARITY.
FT DISULFID 100 116 BY SIMILARITY.
FT DISULFID 111 129 BY SIMILARITY.
FT DISULFID 965 982 BY SIMILARITY.
FT DISULFID 977 995 BY SIMILARITY.
FT MOD_RES 415 415 SULFATION (POTENTIAL).
FT MOD_RES 415 415 SULFATION (POTENTIAL).
FT MOD_RES 424 424 SULFATION (POTENTIAL).
FT MOD_RES 1281 1281 SULFATION (POTENTIAL).
FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 478 478 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 706 706 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 748 748 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 826 826 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 884 884 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 976 976 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 988 988 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1254 1254 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1322 1322 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1363 1363 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1387 1387 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1602 1602 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1671 1671 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1841 1841 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1846 1846 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1856 AA; 209721 MW; 81E7AA0CABABD07D CRC64;

```

```

Query Match 44.58; Score 49; DB 1; Length 1856;
Best Local Similarity 58.88; Pred. No. 7.2;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 4 PNRKTOOYTDQIERPLM 20
DB 1221 PELVTOOYTELIGRPV 1237

```

```

RESULT 4
SECS_YEAST STANDARD; PRT; 1336 AA.
AC P33332:
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
E EXOCYST complex component SEC3 (PSL1 protein).
N SEC3 OR PSL1 OR YER008C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97044444; PubMed=8889515;
RA Hearer B.K., Corbett A., Kweon Y., Petzold A.S., Silver P.,
RA Brown S.S.;
RT "SEC3 mutations are synthetically lethal with profilin mutations and
RL cause defects in diploid-specific bud-site selection."
RL Genetics 144:495-510(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=5288C / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
RA Hyman R., Kayser A., Komp C., Iashkari D., Lew H., Lin D.,
RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
RA Patel F.A., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.

```

```

RN [3]
RP CHARACTERIZATION.
RX MEDLINE=97390810; PubMed=9247645;
RA Finger F.P., Novick P.;
RT "Sec3 is involved in secretion and morphogenesis in Saccharomyces
RL cerevisiae."
RL Mol. Biol. Cell 8:647-662(1997).
RN [4]
RP SEQUENCE OF 793-800.
RX MEDLINE=9713278; PubMed=8978675;
RA Terbush D.R., Maurice T., Roth D., Novick P.;
RT "The Exocyst is a multiprotein complex required for exocytosis in
RL Saccharomyces cerevisiae."
RL EMBO J. 15:6483-6494(1996).
CC -1- FUNCTION: REQUIRED FOR AN EVENT LEADING TO FUSION OF POST-GOLGI
CC VESICLES WITH THE PLASMA MEMBRANE IN YEAST CELLS.
CC -1- SUBUNIT: SEC3, SEC5, SEC6, SEC8, SEC10, SEC15 AND EXO70 ARE
CC COMPONENT OF EXOCYST COMPLEX.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L22204; AAB49380.1; -
DR EMBL; U18778; AAB64541.1; -
DR PIR; S41794; S41794.
DR SGD; S0000810; SEC3.
KW Transport; Protein transport; Golgi stack; Coiled coil.
FT DOMAIN 319 464 COILED COIL (POTENTIAL).
FT DOMAIN 601 621 COILED COIL (POTENTIAL).
FT DOMAIN 1309 1336 COILED COIL (POTENTIAL).
SQ SEQUENCE 1336 AA; 154694 MW; 6F0C1BF13A77985E CRC64;

```

```

Query Match 43.68; Score 48; DB 1; Length 1336;
Best Local Similarity 42.18; Pred. No. 7.3;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
QY 1 TYFPNRKTOOYTDQIERPL 19
DB 498 TYPERSTAHADQLKKPL 516
RESULT 5
VGLG_IHNV STANDARD; PRT; 508 AA.
AC P07923:
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Spike glycoprotein precursor.
GN G.
OS Infectious hematopoietic necrosis virus (strain Round Butte) (IHNV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales.
OC Rhabdoviridae; Nyctirhabdovirus.
OX NCBI_TaxID=11291;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87198856; PubMed=3033264;
RA Koener J.F., Passavant C.W., Kutach G., Leong J.;
RT "Nucleotide sequence of a cDNA clone carrying the glycoprotein gene
RL of infectious hematopoietic necrosis virus, a fish rhabdovirus."
RL J. Virol. 61:1342-1349(1987).
RN [2]
RP REVISIONS.
RA Leong J.;
RL Submitted (XXX-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL, M16023; AAB59926.1; -.
CC PIR: A29532; VGVNFR.
CC InterPro: IPR001903; Rhabd_glycop.
CC InterPro: IPR002417; Spike_prot.
CC Pfam: PF00974; Rhabd_glycop.1.
CC PRINTS: PR00796; SPIKEPROTEIN.
CC Transmembrane; Envelope protein; Glycoprotein; Signal.
CC SIGNAL 1 20
CC CHAIN 21 508 SPIKE GLYCOPROTEIN.
CC FT CARBOHYD 56 56 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (LOCAL).
CC FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 506 506 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 508 AA; 56799 MW; CF65376C626FE89B CRC64;

Query Match 40.0%; Score 44; DB 1; Length 508;
Best Local Similarity 42.1%; Pred. No. 12;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 YFPNKGTOQYTDQIERPLM 20
||| ||| ||| : : :
Db 144 YFPAPRGQWYTDNVQNDLI 162

RESULT 6
CH60_PASMU STANDARD; PRT; 547 AA.
AC 059687;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
GN GROEL OR MOXA OR GROEL OR PM1107.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_Taxid=747;
RA [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PI059;
RC MEDLINE=96105224; PubMed=8529887;
RA Love B.C., Hansen L.M., Hirsch D.C.;
RT "Cloning and sequence of the groEL heat-shock operon of Pasteurella
RT multocida";
RN Gene 166:179-180(1995).
[12]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RC MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whitlam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida PM70";
RT Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CC CONDITIONS (BY SIMILARITY).
CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
CC 7 SUBUNITS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL: U30165; AAB4916.1; -.
CC EMBL: AE006151; AAK03191.1; -.
CC HSSP: P06139; IJON.
CC InterPro: IPR001844; Chaperin_Cpn60.
CC InterPro: IPR002423; Cpn60/TCp-1.
CC Pfam: PF00118; Cpn60_TCp1.1.
CC PRINTS: PR00298; CHAPERONIN60.
CC PROSITE: PS00304; TCOMPLEXTCp1.
CC PROSITE: PS00296; CHAPERONINS_CPN60; 1.
CC Chaperone; ATP-binding; Complete proteome.
CC CONFLICT 424 424 S -> V (IN REF. 1).
CC CONFLICT 464 464 I -> V (IN REF. 1).
CC SEQUENCE 547 AA; 57291 MW; EBB95B9F2358B55E CRC64;

Query Match 40.0%; Score 44; DB 1; Length 547;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 NKGTOQYTDQIERPLM 20
| | | | | : : :
Db 479 NAGTEQYGDVMEIGIL 494

RESULT 7
CH60_HAEIN STANDARD; PRT; 548 AA.
AC P43733;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
GN GROEL OR MOXA OR GROEL OR HI0543.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_Taxid=727;
RA [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KM20 / ATCC 51907;
RC MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uitterback T.R., Hanna M.C., Nguyen D.T., Saudak D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fritchman J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd";
RN Science 269:496-512(1995).
[12]
RP FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
RP PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
RP CONDITIONS (BY SIMILARITY).
CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
CC 7 SUBUNITS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: U32736; AAC22201.1; -.
CC HSSP: P06139; IJON.
CC TIGR: HT0543; -.
CC InterPro: IPR001844; Chaperin_Cpn60.
CC InterPro: IPR002423; Cpn60/TCp-1.


```

CC FT CONFLICT 32 32 G -> A (IN REF. 1).
CC FT CONFLICT 47 47 N -> S (IN REF. 1).
CC FT CONFLICT 57 57 S -> P (IN REF. 1).
CC FT CONFLICT 123 123 Q -> R (IN REF. 1).
CC FT CONFLICT 397 397 R -> N (IN REF. 1).
CC FT CONFLICT 592 592 A -> G (IN REF. 1).
CC FT CONFLICT 679 679 A -> P (IN REF. 1).
CC SQ SEQUENCE 763 AA; 85584 MW; 7C3A52A3CBB04FB5 CRC64;

Query Match 39.5% Score 43.5: DB 1: Length 763;
Best Local Similarity 47.6% Pred. No. 22;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

Db 4 PNRGTQ-----YTDQIERPL 19
||| | | | | | | | | |
46 PMSDQVSALSVDQIQPL 66

RESULT 9
A-RAT
DYR-RAT STANDARD: PRT: 763 AA.
063470:
15-JUL-1999 (Rel. 38, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Dual-specificity tyrosine-phosphorylation regulated kinase 1A
(EC 2.7.1.-) (Protein kinase mindbrain homolog) (MNBH) (RP86) (Dual
specificity YAK1-related kinase).
DYRK1A OR DYRK.
Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A., AND MUTAGENESIS OF TYR-219.
RC STRAIN-Sprague-Dawley; TISSUE=Brain;
MEDLINE=66216443; PubMed=8631952;
RA Kennip H., Becker W., Heukelbach J., Wilmes A., Schuermann A.,
Huppertz G., Kalnulainen H., Joost H.-G.;
"Tyrk, a dual specificity protein kinase with unique structural
features whose activity is dependent on tyrosine residues between
subdomains Vir and VIII."
J. Biol. Chem. 271:3488-3495(1996).
[2]
REVISONS.
RA Kennip H.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
[3]
CHARACTERIZATION.
MEDLINE=98421512; PubMed=9748265;
RA Becker W., Weber Y., Wetzel K., Elmdbler K., Tejedor F.J.,
Joost H.-G.;
"Sequence characteristics, subcellular localization, and substrate
specificity of DYRK-related kinases, a novel family of dual
specificity protein kinases."
J. Biol. Chem. 273:25893-25902(1998).
CC -1- FUNCTION: MAY PLAY A ROLE IN A SIGNALING PATHWAY REGULATING
NUCLEAR FUNCTIONS OF CELL PROLIFERATION. PHOSPHORYLATES SERINES,
THREONINES AND TYROSINES RESIDUES IN ITS SEQUENCE AND IN EXOGENOUS
SUBSTRATES.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: UROGENITAL.
CC -1- PTM: AUTOPHOSPHORYLATED ON TYR RESIDUES (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
MNB/DYRK SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation at
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entitles requires a license agreement (See http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; X79769; CAAS6164.1; -.
DR HSSP; Q00534; 1b1b.
DR InterPro; IPR000719; Euk_Pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR Transfaser; Serine/threonine-protein kinase; Tyrosine-protein kinase;
KW ATP-binding; Nuclear protein; Phosphorylation; Alternative splicing.
FT DOMAIN 117 134
FT FT (POTENTIAL).
FT FT PROTEIN KINASE.
FT FT NP_BIND 159 479
FT FT BINDING 165 173
FT FT BINDING 188 188
FT FT ACT_SITE 287 287
FT FT DOMAIN 509 515
FT FT DOMAIN 599 602
FT FT DOMAIN 607 619
FT FT DOMAIN 656 672
FT FT DOMAIN 664 671
FT FT MOD_RES 219 219
FT FT MOD_RES 319 319
FT FT MOD_RES 321 321
FT FT VASPLIC 70 78
FT FT MUTAGEN 219 219
FT FT
FT FT
SQ SEQUENCE 763 AA; 85541 MW; CB5EC7ECAC1P9A47 CRC64;
Query Match 39.5%; Score 43.5; DB 1; Length 763;
Best Local Similarity 47.6%; Pred. No. 22;
Matches 10; Conservative 3; Indels 5; Gaps 1;
OY 4 PNKGTOQ-----YTDIERPL 19
   |||  |||  |||  |||
DB 46 PNIDQGVSAVSVDIQDQL 66
-----
RESULT 10
CH60_ACTPL STANDARD; PRT; 546 AA.
AC P94166;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
GN GROEL OR MOXA OR GROEL.
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Actinobacillus.
OC NCBI_TaxID=715;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=S 4074 / Serotype 1;
RX MEDLINE=97189570; PubMed=9037757;
RA Verina G., Sirois M., Clatroux N., Boissinot M.;
RT "Cloning and characterization of the groE locus from Actinobacillus
RT pleuropneumoniae.";
RL FEBS Microbiol. Lett. 147:11-16(1997).
CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CC CONDITIONS (BY SIMILARITY).
CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
CC 7 SUBUNITS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U55016; AAB51437.1; -
 DR HSP: P06139; 1GRL.
 DR InterPro: IPR001844; Chaprin Cpn60.
 DR InterPro: IPR002423; Cpn60/TCF-1.
 DR Pfam: PF01118; Cpn60_TCF1.1.
 DR PRINTS: PR00298; CHAPERONIN60.
 DR PRINTS: PR00304; TCOMPLEXCP1.
 DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.
 DR Chaperone; ATP-binding.
 FT INIT_MEN 0
 FT SEQUENCE 546 AA; 57513 MW; E84BB72C9BD3DB56 CRC64;

Query Match 39.1%; Score 43; DB 1; Length 546;
 Best Local Similarity 42.1%; Pred. No. 18;
 Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 YFPNGTQYTDIERP 20
 Db 476 YGYNAGTEGYCDMLEMGCIL 494

RESULT 11
 ID 7GAC_BREDI STANDARD; PRT; 720 AA.
 AC 09L5D6;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Glutaryl-7-aminocephalosporanic acid acylase precursor (EC 3.5.1.-)
 DE (Glutaryl-7-ACA-acylase).
 OS Brevundimonas diminuta (Pseudomonas diminuta).
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Brevundimonas.
 OX NCBI_TaxID=293;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN-KAC-1;

RA Kim D.-W., Kang S.-M., Yoon K.-H.;
 RT "Cloning and the nucleotide sequence of a Pseudomonas diminuta KAC-1
 RT glutaryl-7-aminocephalosporanic acid acylase gene."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 -1- PATHWAY: SECOND STEP IN THE BIOCONVERSION OF CEPHALOSPORIN TO 7-
 ACA.
 -1- SUBCELLULAR LOCATION: Periplasmic (Potential).
 -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY 545.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF251710; AAF64242.1; -
 DR MEROPS: S45.002; -
 DR InterPro: IPR002692; Penicill_amidase.
 DR Pfam: PF01804; Penicill_amidase; 1.
 KW Hydrolyase; Periplasmic; Antibiotic resistance; Zymogen; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 720 POTENTIAL.
 FT GLUTARYL-7-AMINOCEPHALOSPORANIC ACID
 FT ACT_SITE 290 290 ACYLASE.
 FT SEQUENCE 720 AA; 79779 MW; AD624797845CC39B CRC64;

Query Match 39.1%; Score 43; DB 1; Length 720;

Best Local Similarity 72.7%; Pred. No. 25;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 GTQGYTDIER 17
 Db 680 GTTHYSDIER 690

RESULT 12
 ID UBPO_HUMAN STANDARD; PRT; 913 AA.
 AC 09BX07;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ubiquitin carboxyl-terminal hydrolase 26 (EC 3.1.2.15) (Ubiquitin
 DE thiolesterase 26) (Ubiquitin-specific processing protease 26)
 DE (Deubiquitinating enzyme 26).
 GN USP26.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN 11
 RP SEQUENCE FROM N.A.
 RC TISSUE-Testis;

RX MEDLINE=21175748; PubMed=11279525;
 RA Wang P.-Y., McCarrey J.R., Yang F., Page D.C.;
 RT "An abundance of X-linked genes expressed in spermatogonia."
 RL Nat. Genet. 27:422-426(2001).
 CC -1- FUNCTION: INVOLVED IN THE UBIQUITIN-DEPENDENT PROTEOLYTIC PATHWAY
 CC IN CONJUNCTION WITH THE 26S PROTEASOME (BY SIMILARITY)

CC -1- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
 CC ubiquitin + a thiol.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C19.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF285593; AAK31972.1; -
 DR MEROPS: C19.046; -
 DR Genew: HGNC:13485; USP26.
 DR MIM: 300309; -
 DR InterPro: IPR001394; UCH-2.
 DR Pfam: PF00443; UCH-2; 1.
 DR PROSITE: PS00972; UCH-2; 1.
 DR PROSITE: PS00973; UCH-2; 1.
 DR PROSITE: PS50235; UCH-2; 3; 1.
 KW Ub1 conjugation pathway; Hydrolase; Thiol protease; Multigene family.
 FT ACT_SITE 304 304 BY SIMILARITY.
 FT ACT_SITE 832 832 BY SIMILARITY.
 FT ACT_SITE 841 841 BY SIMILARITY.
 FT SEQUENCE 913 AA; 104047 MW; BB70B09CDBA3C48 CRC64;

Query Match 39.1%; Score 43; DB 1; Length 913;
 Best Local Similarity 50.0%; Pred. No. 33;
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 FPKNGTQYTDIERP 18
 Db 760 FPKNGTQYTDIERP 775

RESULT 13
 ID UBIG_RHIME STANDARD; PRT; 248 AA.
 AC 092MK1;


```

RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kamehisa M., Yanashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001) .
RN [2]
RP SEQUENCE FROM N.A.
RC SPRAIN-WM2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takuchih F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002) .
CC -1- FUNCTION: Provides the (R)-glutamic acid required for cell wall
CC biosynthesis (By similarity).
CC -1- CATALYTIC ACTIVITY: L-glutamate -> D-glutamate.
CC -1- PATHWAY: Peptidoglycan biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE ASPARTATE/Glutamate RACEMASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AP003361; BAB57313.1; -
DR EMBL, AP003132; BAB42246.1; -
DR EMBL, AP004825; BAB94898.1; -
DR HSSP, P56868, 1874.
DR InterPro: IPR001920; Asp/Glu_race.
DR InterPro: IPR004391; Glu_race.
DR Pfam, PF01117; Asp_Glu_race; 1.
DR TIGRFAMS: TIGR00067; glut_race; 1.
DR PROSITE: PS00923; ASP_GLU_RACEMASE_1; 1.
DR PROSITE: PS00924; ASP_GLU_RACEMASE_2; 1.
KW Peptidoglycan-synthesis; Cell wall; Isomerase; Complete proteome.
SQ SEQUENCE 266 AA; 29698 MW; 93BC35DBA42C7B0E CMC64;

```

Search completed: December 11, 2002, 15:26:27
Job time : 11 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 15:25:17 ; Search time 15 Seconds

(without alignments)
128.179 Million cell updates/sec

Title: US-09-991-433-48

Perfect score: 110

Sequence: 1 TYFPNKGTOQYTDQIERPLM 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database : PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	781	1	VCPV19
2	58	52.7	458	2	B90625
3	55	50.0	459	2	T11333
4	54	49.1	458	2	B90623
5	53	48.2	458	2	B90615
6	53	48.2	459	2	T11202
7	52	47.3	455	2	B90619
8	52	47.3	459	2	T11084
9	51	46.4	458	2	T11528
10	51	46.4	458	2	T11176
11	51	46.4	458	2	B99613
12	50	45.5	458	2	B90637
13	49	44.5	458	2	B90621
14	49	44.5	13055	2	T16580
15	48	43.6	1336	2	S41794
16	46	41.8	458	2	B90617
17	46	41.8	459	2	T11031
18	45	41.8	714	2	S76082
19	45	41.4	519	2	T45768
20	45	40.9	299	2	AG3274
21	45	40.9	868	2	T06827
22	44	40.0	71	2	D87036
23	44	40.0	508	1	VGWNR
24	44	40.0	547	2	JC4519
25	44	40.0	548	2	CC4076
26	44	40.0	1173	2	B97208
27	43.5	39.5	754	2	JC4898
28	43	39.1	89	2	I68530
29	43	39.1	177	2	AH0069

30	43	39.1	459	2	T11189	NADH2 dehydrogenas
31	43	39.1	1577	2	T15851	hypothetical prote
32	42	38.2	88	2	A44929	protein-tyrosine-p
33	42	38.2	333	2	D88970	protein f15E11.11
34	42	38.2	398	2	I65440	protein-tyrosine-p
35	42	38.2	514	2	B86164	hypothetical prote
36	42	38.2	538	2	A85075	probable transposo
37	42	38.2	669	2	S14535	asparagine-rich pr
38	42	38.2	765	2	A00778	beta-glucosidase (
39	42	38.2	859	2	T06429	lipoxigenase (EC 1
40	42	38.2	1262	1	B48758	protein-tyrosine-p
41	42	38.2	1496	1	A48758	protein-tyrosine-p
42	42	38.2	1501	2	I58148	protein-tyrosine-p
43	42	38.2	1863	2	S46217	protein-tyrosine-p
44	42	38.2	1907	2	S50893	protein-tyrosine-p
45	41	37.3	206	2	A81231	50S ribosomal prot

ALIGNMENTS

RESULT 1

VCPV19

coat protein VP1 - parvovirus B19 (strain Au)

C:Species: parvovirus B19

A:Note: host Homo sapiens (man)

C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999

C:Accession: A24299

R:Shade, R.O.; Blundell, M.C.; Cotmore, S.F.; Tattersall, P.; Astell, C.R.

J. Virol. 58, 921-936, 1986

A:Title: Nucleotide sequence and genome organization of human parvovirus B19 isolated

A:Reference number: A24299; MUID:86200451; PMID:3701931

A:Accession: A24299

A:Molecule type: DNA

A:Residues: 1-781 <SHA>

A:Cross-references: EMBL:M13178; NID:Q333375; PIDN:AA66867.1; PID:Q333377

C:Superfamily: parvovirus coat protein

C:Keywords: coat protein

Query Match

Best Local Similarity 100.0%; Score 110; DB 1; Length 781;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQIERPLM 20

DB 618 TYFPNKGTOQYTDQIERPLM 637

RESULT 2

B90625

NADH dehydrogenase chain 4 ND4 [imported] - Casuarinus casuarinus mitochondrion

C:Species: mitochondrion Casuarinus casuarinus

C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 03-Aug-2001

C:Accession: B90625

R:Haddarath, O.; Baker, A.J.

Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001

A:Title: Complete mitochondrial DNA genome sequences of extinct birds: ratite pnyl

A:Reference number: B99613; MUID:21263106; PMID:11370967

A:Accession: B90625

A:Molecule type: DNA

A:Residues: 1-458 <KUR>

A:Cross-references: GB:NC-002778; NID:Q14141873; PIDN:NP-115359.1; GSPDB:GN00162

C:Genetics:

A:Gene: ND4

A:Genome: mitochondrion

A:Superfamily: NADH dehydrogenase (ubiquinone) chain 4

C:Keywords: mitochondrion

Query Match

Best Local Similarity 54.5%; Score 58; DB 2; Length 458;

Matches 12; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

OY 1 TYFPNKGTOOYT--DQIERPLM 20
||||| 1:1 ||| 11:
Db 45 TYFPHKNLQWMTGIDQISAPLL 66

RESULT 3

T11333
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - Corvus frugilegus mitochondrion
C:Species: mitochondrion Corvus frugilegus
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
C:Accession: T11333
R:Haddad, A.; Arnason, U.
Proc. R. Soc. Lond. B Biol. Sci. 266, 305-309, 1999
A:Title: Analyses of mitochondrial DNA nest ratite birds within the Neognathae-supporting
A:Reference number: 217262
A:Accession: T11333
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-459 <HAR>
C:Cross-references: EMBL:Y18522; PIDN:CAA7204.1
C:Genetics:
A:Genome: mitochondrion
A:Note: NADH4
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 50.0%; Score 55; DB 2; Length 459;
Best Local Similarity 50.0%; Pred. No. 0.6;
Matches 11; Conservative 4; Mismatches 5; Indels 2; Gaps 1;
-OY 1 TYFPNKGTOOYT--DQIERPLM 20
||||| 1:1 ||| 11:
Db 45 TYPSKGLTHTSIDQISSPLL 66

-RESULT 4

B90623
NADH dehydrogenase chain 4 ND4 [imported] - Anomalopteryx didiformis mitochondrion
C:Species: mitochondrion Anomalopteryx didiformis
C>Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 03-Aug-2001
C:Accession: B90623
R:Haddad, O.; Baker, A.J.
Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001
A:Title: Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogene
A:Reference number: A99613; MUID:21263106; PMID:11370967
A:Accession: B90623
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-458 <KUR>
C:Cross-references: GB:NC_002779; NID:g14141923; PIDN:NP_115372.1; GSPDB:GN00161
C:Genetics:
A:Gene: ND4
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C:Keywords: mitochondrion

Query Match 49.1%; Score 54; DB 2; Length 458;
Best Local Similarity 54.5%; Pred. No. 0.87;
Matches 12; Conservative 2; Mismatches 6; Indels 2; Gaps 1;
OY 1 TYFPNKGTOOYT--DQIERPLM 20
||||| 1:1 ||| 11:
Db 45 TYFPYKNLTQWMTGIDQISAPLL 66

RESULT 5

B90615
NADH dehydrogenase chain 4 ND4 [imported] - Dromaius novaehollandiae mitochondrion
C:Species: mitochondrion Dromaius novaehollandiae
C>Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 03-Aug-2001
C:Accession: B90615

R:Haddad, O.; Baker, A.J.
Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001
A:Title: Complete mitochondrial DNA genome sequences of extinct birds: ratite phylog
A:Reference number: A99613; MUID:21263106; PMID:11370967
A:Accession: B90615

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-458 <KUR>
C:Cross-references: GB:NC_002784; NID:g14141887; PIDN:NP_115437.1; GSPDB:GN00157
C:Genetics:
A:Gene: ND4
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C:Keywords: mitochondrion

Query Match 48.2%; Score 53; DB 2; Length 458;
Best Local Similarity 50.0%; Pred. No. 1.3;
Matches 11; Conservative 3; Mismatches 6; Indels 2; Gaps 1;

OY 1 TYFPNKGTOOYT--DQIERPLM 20
||||| 1:1 ||| 11:
Db 45 TYFPHKNLQWMTGIDQISAPLL 66

RESULT 6

T11202
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - Vidua chalybeata mitochondrio
C:Species: mitochondrion Vidua chalybeata
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
C:Accession: T11202
R:Mindell, D.P.; Sorenson, M.D.; Dimcheff, D.E.
Proc. Natl. Acad. Sci. U.S.A. 95, 10693-10697, 1998
A:Title: Multiple independent origins of mitochondrial gene order in birds.
A:Reference number: 217242

A:Accession: T11202
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-459 <MIN>
C:Cross-references: EMBL:AF090341; NID:g4894501; PID:g4894511; PIDN:AD32534.1
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation

Query Match 48.2%; Score 53; DB 2; Length 459;
Best Local Similarity 50.0%; Pred. No. 1.3;
Matches 11; Conservative 3; Mismatches 6; Indels 2; Gaps 1;

OY 1 TYFPNKGTOOYT--DQIERPLM 20
||||| 1:1 ||| 11:
Db 45 TYPSKGLTHTSIDQISSPLL 66

RESULT 7

B90619
NADH dehydrogenase chain 4 ND4 [imported] - Apteryx haastii mitochondrion
C:Species: mitochondrion Apteryx haastii
C>Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 03-Aug-2001
C:Accession: B90619
R:Haddad, O.; Baker, A.J.
Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001
A:Title: Complete mitochondrial DNA genome sequences of extinct birds: ratite phylog
A:Reference number: A99613; MUID:21263106; PMID:11370967
A:Accession: B90619

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-455 <KUR>
C:Cross-references: GB:NC_002782; NID:g14141943; PIDN:NP_115411.1; GSPDB:GN00159
C:Genetics:
A:Gene: ND4
A:Genome: mitochondrion

A:Genetic code: SGCI
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C:Keywords: mitochondrion

Query Match 47.3%; Score 52; DB 2; Length 455;
Best Local Similarity 54.5%; Pred. No. 1.8;
Matches 12; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

OY 1 TYPFNKGTQOYT--DQIERPLM 20
||||| :|:| ||| ||:
Db 45 TYPFKYLSQWTGIDQISAPLL 66

RESULT 8

T11084
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - Falco peregrinus mitochondrion

C:Species: mitochondrion Falco peregrinus
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
C:Accession: T11084
R:Mindell, D.P.; Sorenson, M.D.; Dimcheff, D.E.
Proc. Natl. Acad. Sci. U.S.A. 95, 10693-10697, 1998
A:Title: Multiple independent origins of mitochondrial gene order in birds.
A:Reference number: Z17242
A:Accession: T11084
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-459 <MIN>
A:Cross-references: EMBL:AF090338; NID:g4894462; PID:g4894472; PIDN:AAD32501.1

C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGCI
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 47.3%; Score 52; DB 2; Length 459;
Best Local Similarity 50.0%; Pred. No. 1.8;
Matches 11; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

OY 1 TYPFNKGTQOYT--DQIERPLM 20
||||| :|:| ||| ||:
Db 45 TYPFKYLSQWTGIDQISAPLL 66

RESULT 9

T11528
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - ostrich mitochondrion

C:Species: mitochondrion Struthio camelus (ostrich)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
C:Accession: T11528
R:Harild, A.
submitted to the EMBL Data Library, September 1998
A:Reference number: Z17278
A:Accession: T11528
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-458 <HAR>
A:Cross-references: EMBL:Y12025; PIDN:CAA72753.1

C:Genetics:
A:Gene: NADH4
A:Genome: mitochondrion
A:Genetic code: SGCI
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 46.4%; Score 51; DB 2; Length 458;
Best Local Similarity 50.0%; Pred. No. 2.7;
Matches 11; Conservative 3; Mismatches 6; Indels 2; Gaps 1;

OY 1 TYPFNKGTQOYT--DQIERPLM 20
||||| :|:| ||| ||:
Db 45 TYPFKYLSQWTGIDQISAPLL 66

RESULT 10

T11176

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - greater rhea mitochondrion

C:Species: mitochondrion Rhea americana (greater rhea, common rhea)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
C:Accession: T11176
R:Mindell, D.P.; Sorenson, M.D.; Dimcheff, D.E.
Proc. Natl. Acad. Sci. U.S.A. 95, 10693-10697, 1998
A:Title: Multiple independent origins of mitochondrial gene order in birds.
A:Reference number: Z17242
A:Accession: T11176
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-458 <MIN>
A:Cross-references: EMBL:AF090339; NID:g4894475; PID:g4894485; PIDN:AAD32512.1

R:Harild, A.
submitted to the EMBL Data Library, May 1999
A:Reference number: Z17270
A:Accession: T11424
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-62, 'P', 64-181, 'I', 183-268, 'V', 270-456, 'A', 458 <HAR>
A:Cross-references: EMBL:Y16884; PIDN:CAA76509.1

C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGCI
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 46.4%; Score 51; DB 2; Length 458;
Best Local Similarity 50.0%; Pred. No. 2.7;
Matches 11; Conservative 3; Mismatches 6; Indels 2; Gaps 1;

OY 1 TYPFNKGTQOYT--DQIERPLM 20
||||| :|:| ||| ||:
Db 45 TYPFKYLSQWTGIDQISAPLL 66

RESULT 11

B9613

NADH dehydrogenase chain 4 ND4 [imported] - Struthio camelus mitochondrion

C:Species: mitochondrion Struthio camelus
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 03-Aug-2001
C:Accession: B9613
R:Haddad, O.; Baker, A.J.
Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001
A:Title: Complete mitochondrial DNA genome sequences of extinct birds: ratite phylog
A:Reference number: A9613; MUID:21263106; PMID:11370967
A:Accession: B9613
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-458 <KUR>
A:Cross-references: GB:NC_002785; NID:g14141915; PIDN:NP_115450.1; GSPDB:GN00156

C:Genetics:
A:Gene: ND4
A:Genome: mitochondrion
A:Genetic code: SGCI
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C:Keywords: mitochondrion

Query Match 46.4%; Score 51; DB 2; Length 458;
Best Local Similarity 50.0%; Pred. No. 2.7;
Matches 11; Conservative 3; Mismatches 6; Indels 2; Gaps 1;

OY 1 TYPFNKGTQOYT--DQIERPLM 20
||||| :|:| ||| ||:
Db 45 TYPFKYLSQWTGIDQISAPLL 66

RESULT 12

B90627

NADH dehydrogenase chain 4 ND4 [imported] - Eudromia elegans mitochondrion

C:Species: mitochondrion Eudromia elegans

C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 03-Aug-2001
C:Accession: B90627
R:Haddrath, O.; Baker, A.J.
Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001
A:Title: Complete mitochondrial DNA genome sequences of extinct birds: rattle phylogenetic
A:Reference number: A99613; MUID:21263106; PMID:11370967
A:Accession: B90627
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1458 <KUR>
A:Cross-references: GB:NC_002772; NID:g14141810; PIDN:NP_115282.1; GSPDB:GN00163
C:Genetics:
A:Gene: ND4
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C:Keywords: mitochondrion

RESULT 13

B90621

NADH dehydrogenase chain 4 ND4 [Imported] - *Tinamus major* mitochondrion

C:Species: mitochondrion *Tinamus major*

C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 03-Aug-2001

C:Accession: B90621

R:Haddrath, O.; Baker, A.J.

Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001

A:Title: Complete mitochondrial DNA genome sequences of extinct birds: rattle phylogenetic

A:Reference number: A99613; MUID:21263106; PMID:11370967

A:Accession: B90621

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-458 <KUR>

A:Cross-references: GB:NC_002781; NID:g14141957; PIDN:NP_115398.1; GSPDB:GN00160

C:Genetics:

A:Gene: ND4

A:Genome: mitochondrion

A:Genetic code: SSCI

C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4

Keywords: mitochondrion

Query Match 44.5%; Score 49; DB 2; Length 458;

Best Local Similarity 45.5%; Pred. No. 5.6;

Matches 10; Conservative 6; Mismatches 4; Indels 2; Gaps 1;

OY 1 TYFPNKGTQOQT--DQIERPLM 20

:|||:|:|:|:|:|:|

Db 45 SYFPKFTSSQWSSIDQISAPLL 66

RESULT 14

T16580

hypothetical protein K07E12.1 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T16580

R:Pulton, L.

submitted to the EMBL Data Library, May 1994

A:Description: The sequence of *C. elegans* cosmid K07E12.

A:Reference number: Z18340

A:Accession: T16580

A:Status: Preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-13055 <FUL>

A:Cross-references: EMBL:U00054; NID:g485140; PID:g485141; PIDN:AAA50715.1; CESP:K07E12

```

A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:K07E12.1
A:Introns: 46/3; 85/2; 201/1; 278/2; 470/2; 817/1; 927/1; 960/3; 1265/1; 1322/1; 1478/1; 6014/3; 6159/3; 6665/2; 7266/3; 7895/3; 8669/3; 8726/3; 9803/3; 10937/3; 12234/2; 1
Query Match          44.5%  Score 49;  DB 2;  Length 13055;
Best Local Similarity 60.0%;  Pred. No. 2.3e+02;
Matches          9;  Conservative 2;  Mismatches 4;  Indels 0;  Gaps 0;
OY      4  PNKGTOQYTDQIERP 18
      || : |||| 1:1
Db      1879  PNWETIQYTDDEAP 1893

RESULT 15
S41794
SEC3 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YER008c; PSL1 protein
C:Species: Saccharomyces cerevisiae
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text-change 21-Jul-2000
C:Accession: S41794; S50466; S72238
R:Haarer, B.K.; Petzold, A.S.; Brown, S.S.
Submitted to the EMBL Data Library, July 1993
A:Description: Identification of mutations that are synthetically lethal with altered
A:Reference number: S41793
A:Accession: S41794
A:Molecule type: DNA
A:Residues: 1-1336 <HAA>
A:Cross-references: EMBL:L22204; NID:g347714; PID:g347716
R:Dietrich, F.S.
Submitted to the EMBL Data Library, December 1994
A:Description: The sequence of S. cerevisiae cosmid 9537, 9581, 9495, 9867, and lambd
A:Reference number: S50459
A:Accession: S50456
A:Molecule type: DNA
A:Residues: 1-1336 <DIE>
A:Cross-references: EMBL:U18778; NID:g603592; PID:g603600; MIPS:YER008c
R:Haarer, B.K.; Corbett, A.; Kweon, Y.; Petzold, A.S.; Silver, P.; Brown, S.S.
Genetics 144, 495-510, 1996
A:Title: SEC3 mutations are synthetically lethal with profilin mutations and cause de
A:Reference number: S72237; MUID:97044444; PMID:8889515
A:Accession: S72238
A:Molecule type: DNA
A:Residues: 1-1336 <HAW>
A:Cross-references: EMBL:L22204; NID:g347714; PID:AB49380.1; PID:g347716
C:Genetics:
A:Gene: SGD:SEC3; PSL1
A:Cross-references: MIPS:YER008c; SGD:S0000810
A:Map position: 5R

Query Match          43.6%  Score 48;  DB 2;  Length 1336;
Best Local Similarity 42.1%;  Pred. No. 27;
Matches          8;  Conservative 5;  Mismatches 6;  Indels 0;  Gaps 0;
OY      1  TYFPNKGTOQYTDQIERP 19
      || | : | : ||::||
Db      498  TYPERSTARAHDLQKPL 516

Search completed: December 11, 2002, 15:27:26
Job time : 17 secs

```

```

RESULT 15
S41794
SEC3 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: Protein YER008c; PSL1 protein
C:Species: Saccharomyces cerevisiae
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 21-Jul-2000
C:Accession: S41794; S50466; S72238
R:Haarer, B.K.; Petzold, A.S.; Brown, S.S.
Submitted to the EMBL Data Library, July 1993
A:Description: Identification of mutations that are synthetically lethal with altered
A:Reference number: S41794
A:Accession: S41794
A:Residues: 1-1336 <HAA>
A:Molecule type: DNA
A:Cross-references: EMBL:L22204; NID:g347714; PID:g347716
R:Dietrich, F.S.
Submitted to the EMBL Data Library, December 1994
A:Description: The sequence of S. cerevisiae cosmid 9537, 9581, 9495, 9867, and lambd
A:Reference number: S50459
A:Accession: S50466
A:Molecule type: DNA
A:Residues: 1-1336 <DIE>
A:Cross-references: EMBL:U18778; NID:g603592; PID:g603600; MIPS:YER008c
R:Haarer, B.K.; Corbett, A.; Kweon, Y.; Petzold, A.S.; Silver, P.; Brown, S.S.
Genetics 144, 495-510, 1996
A:Title: SEC3 mutations are synthetically lethal with profilin mutations and cause del
A:Reference number: S72237; MUID:9704444; PMID:8889515
A:Accession: S72238
A:Molecule type: DNA
A:Residues: 1-1336 <HAM>
A:Cross-references: EMBL:L22204; NID:g347714; PID:NAAA49380.1; PID:g347716
C:Genetics:
A:Gene: SGD:SEC3; PSL1
A:Cross-references: MIPS:YER008c; SGD:S0000810
A:Map position: 5R

Query Match 43.6%; Score 48; DB 2; Length 1336;
Best Local Similarity 42.1%; Pred. No. 27;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 TYFNNKGTQQTDDIERPL 19
|||::|::|::|::|::|
Db 498 TYPERSTARAHDDLRPL 516

Search completed: December 11, 2002, 15:27:26
Job time : 17 secs

```

Search completed: December 11, 2002, 15:27:26
Job time : 17 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 15:26:17 ; Search time 10 Seconds
(without alignments)
32.485 Million cell updates/sec

Title: US-09-991-433-48
Perfect score: 110
Sequence: 1 TYPNKGTOGYTDQIERPLM 20

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEM_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCIT_NEM_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEM_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEM_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCITUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US09_NEM_PUB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEM_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEM_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43.5	39.5	50	10 US-09-864-761-34120	Sequence 34120, A
2	43	39.1	913	10 US-09-801-574-65	Sequence 65, Appl
3	42	38.2	1502	9 US-09-808-602-54	Sequence 54, Appl
4	42	38.2	1948	9 US-09-808-602-54	Sequence 55, Appl
5	41	37.3	266	10 US-09-815-242-5346	Sequence 5346, Ap
6	41	37.3	266	10 US-09-815-242-12340	Sequence 12340, A
7	41	37.3	266	10 US-09-815-242-12989	Sequence 12989, A
8	41	37.3	283	10 US-09-815-242-13476	Sequence 13476, A
9	40	36.4	436	10 US-09-815-242-13218	Sequence 13218, A
10	40	36.4	557	12 US-10-052-586-326	Sequence 326, App
11	40	36.4	758	9 US-10-117-846-16	Sequence 16, Appl
12	40	36.4	1244	10 US-09-815-915-8	Sequence 8, Appl
13	39	35.5	40	10 US-09-864-761-40422	Sequence 40422, A
14	39	35.5	428	10 US-09-347-331-10	Sequence 10, Appl
15	38.5	35.0	232	10 US-09-815-242-13601	Sequence 13601, A
16	38	34.5	426	10 US-09-925-300-1502	Sequence 1502, Ap
17	38	34.5	472	10 US-09-745-763-36	Sequence 36, Appl
18	37	33.6	130	9 US-09-764-868-803	Sequence 803, App
19	37	33.6	295	10 US-09-815-242-11833	Sequence 11833, A

20	37	33.6	324	12 US-10-062-254-364	Sequence 364, App
21	37	33.6	380	9 US-10-020-674-2	Sequence 2, Appl
22	37	33.6	405	10 US-09-864-761-38102	Sequence 38102, A
23	37	33.6	414	9 US-10-108-605-87	Sequence 87, Appl
24	37	33.6	433	10 US-09-893-737-106	Sequence 106, App
25	37	33.6	446	10 US-09-859-888-4	Sequence 4, Appl
26	37	33.6	446	10 US-09-925-300-1554	Sequence 1554, Ap
27	37	33.6	467	10 US-09-991-2124-3	Sequence 3, Appl
28	37	33.6	478	9 US-09-859-888-2	Sequence 2, Appl
29	37	33.6	757	9 US-10-117-846-10	Sequence 10, Appl
30	37	33.6	1091	10 US-09-875-423-4	Sequence 4, Appl
31	36.5	33.2	497	9 US-09-866-570A-65	Sequence 65, Appl
32	36.5	33.2	497	10 US-09-866-572A-65	Sequence 65, Appl
33	36	32.7	62	10 US-09-864-761-47146	Sequence 47146, A
34	36	32.7	85	10 US-09-864-761-39996	Sequence 39996, A
35	36	32.7	168	10 US-09-922-217-198	Sequence 198, App
36	36	32.7	168	10 US-09-833-263-198	Sequence 198, App
37	36	32.7	189	9 US-09-792-251-14	Sequence 14, Appl
38	36	32.7	254	10 US-09-815-242-5434	Sequence 5434, Ap
39	36	32.7	254	10 US-09-815-242-12274	Sequence 12274, A
40	36	32.7	425	10 US-09-925-301-1112	Sequence 1112, Ap
41	36	32.7	448	10 US-09-815-242-11362	Sequence 11362, A
42	36	32.7	478	9 US-10-193-295-2	Sequence 2, Appl
43	36	32.7	518	9 US-10-193-295-5	Sequence 5, Appl
44	36	32.7	520	9 US-10-193-295-4	Sequence 4, Appl
45	36	32.7	545	10 US-09-978-242-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-864-761-34120
Sequence 34120, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn. Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
SEQ ID NO 34120
LENGTH: 50
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AP000154.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.8
OTHER INFORMATION: SWISSPROT HIT: Q13627, EVALU 9.00e-24
OTHER INFORMATION: EST_HUMAN HIT: BE817813.1, EVALU 1.00e-22
US-09-864-761-34120

Query Match 39.5%; Score 43.5; DB 10; Length 50;
Best Local Similarity 47.6%; Pred. No. 0.91;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

OY 4 PNRGTQ-----YTDQIERPL 19
||| 1:1111:11
Db 27 PNRDQVSAISYSDIQOPL 47

RESULT 2
US-09-801-574-65
Sequence 65, Application US/09801574
Patent No. US20020081592A1
GENERAL INFORMATION:
APPLICANT: Wang, Peijiang Jeremy
APPLICANT: Page, David C.
TITLE OF INVENTION: Reproduction-Specific Genes
FILE REFERENCE: 0399.2007-002
CURRENT APPLICATION NUMBER: US/09/801,574
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: 60/187,518
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: 60/261,557
PRIOR FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 90
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 65
LENGTH: 913
TYPE: PRT
ORGANISM: Homo sapiens
US-09-801-574-65

Query Match 39.1%; Score 43; DB 10; Length 913;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 3 PNRGTQYTDQIERP 18
||| 1:1:1
Db 760 PNRGTQYTDQIERP 775

RESULT 3
US-09-808-602-54

Sequence 54, Application US/09808602
Patent No. US20020155115A1
GENERAL INFORMATION:
APPLICANT: Vernet, Corine A
APPLICANT: Fernandes, Elma
APPLICANT: Shimkels, Richard A
APPLICANT: Herrman, John L
APPLICANT: Majumder, Kumud
APPLICANT: Mishra, Vishnu
APPLICANT: Mezes, Peter S
APPLICANT: MacDougall, John
TITLE OF INVENTION: NO. US20020155115A1 Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 15966-697 CIP
CURRENT APPLICATION NUMBER: US/09/808,602
CURRENT FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 09/800,198
PRIOR FILING DATE: 2001-03-05
PRIOR APPLICATION NUMBER: 60/186,596
PRIOR FILING DATE: 2000-03-03
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 54
LENGTH: 1502
TYPE: PRT
ORGANISM: Homo sapiens
US-09-808-602-54

Query Match 38.2%; Score 42; DB 9; Length 1502;
Best Local Similarity 60.0%; Pred. No. 76;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 YPNRGTQY 11
|:111:11:1
Db 1058 YPNRGTETY 1067

RESULT 4
US-09-808-602-55
Sequence 55, Application US/09808602
Patent No. US20020155115A1
GENERAL INFORMATION:
APPLICANT: Vernet, Corine A
APPLICANT: Fernandes, Elma
APPLICANT: Shimkels, Richard A
APPLICANT: Herrman, John L
APPLICANT: Majumder, Kumud
APPLICANT: Mishra, Vishnu
APPLICANT: Mezes, Peter S
APPLICANT: MacDougall, John
TITLE OF INVENTION: NO. US20020155115A1 Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 15966-697 CIP
CURRENT APPLICATION NUMBER: US/09/808,602
CURRENT FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 09/800,198
PRIOR FILING DATE: 2001-03-05
PRIOR APPLICATION NUMBER: 60/186,596
PRIOR FILING DATE: 2000-03-03
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 55
LENGTH: 1948
TYPE: PRT
ORGANISM: Homo sapiens
US-09-808-602-55

Query Match 38.2%; Score 42; DB 9; Length 1948;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 YPNRGTQY 11
|:111:11:1
Db 1504 YPNRGTETY 1513

RESULT 5
US-09-815-242-5346
; Sequence 5346, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlson, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5346
; LENGTH: 266
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-09-815-242-5346

Query Match
Best Local Similarity 37.3%; Score 41; DB 10; Length 266;
Matches 11; Conservative 1; Mismatches 7; Indels 2; Gaps 1;

QY 2 YFPNKGTO--QYTDQIERPLM 20
| | | | | : | | | |
Db 41 YGPRGEGVKQYVEIARKLM 61

RESULT 6
US-09-815-242-12340
; Sequence 12340, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlson, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12340
; LENGTH: 266
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-09-815-242-12340

Query Match
Best Local Similarity 37.3%; Score 41; DB 10; Length 266;
Matches 11; Conservative 1; Mismatches 7; Indels 2; Gaps 1;

QY 2 YFPNKGTO--QYTDQIERPLM 20
| | | | | : | | | |
Db 41 YGPRGEGVKQYVEIARKLM 61

RESULT 7
US-09-815-242-12989
; Sequence 12989, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlson, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12989
; LENGTH: 266
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-09-815-242-12989

Query Match
Best Local Similarity 37.3%; Score 41; DB 10; Length 266;
Matches 11; Conservative 1; Mismatches 7; Indels 2; Gaps 1;

QY 2 YFPNKGTO--QYTDQIERPLM 20
| | | | | : | | | |
Db 41 YGPRGEGVKQYVEIARKLM 61

```
RESULT 8
US-09-815-242-13476
; Sequence 13476, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13476
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13476

Query Match          37.3%; Score 41; DB 10; Length 283;
Best Local Similarity 46.7%; Pred. No. 17;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 TYPNKGTOOYTDOI 15
| |||:||||:
205 TSVNKNKSQYTTETV 219

RESULT 9
US-09-815-242-13218
; Sequence 13218, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
```

```
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13218
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13218

Query Match          36.4%; Score 40; DB 10; Length 436;
Best Local Similarity 43.8%; Pred. No. 40;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 4 PNKGTOOYTDOIIRPL 19
| |||:|:|:|
Db 355 PGKGEYMSKLLRL 370

RESULT 10
US-10-052-586-326
; Sequence 326, Application US/10052586
; Patent No. US20020127584A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/052,586
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063564
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063734
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063870
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/064103
```


[illegible]

;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089908

Query Match 36.4%; Score 40; DB 12; Length 557;
Best Local Similarity 50.0%; Pred. No. 52;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 5 NKGTQYTDQIERPLM 20
DB 414 NKATQFTNOTDIPNM 429

US-10-117-846-16
;; Sequence 16, Application US/10117846
;; Patent No. US20020168673A1
;; GENERAL INFORMATION:
;; APPLICANT: Fuller, Margaret T
;; APPLICANT: Hales, Karen G.
;; APPLICANT: Santel, Ansgar H.
;; TITLE OF INVENTION: Mitofusins, Pzo Homologs and Functional
;; FILE REFERENCE: STAN-063CIP3
;; CURRENT APPLICATION NUMBER: US/10/117,846
;; PRIOR FILING DATE: 2002-04-05
;; PRIOR APPLICATION NUMBER: 09/413,285
;; PRIOR FILING DATE: 1999-10-06
;; PRIOR APPLICATION NUMBER: PCT/US00/27871
;; PRIOR FILING DATE: 2000-10-06
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 16
;; LENGTH: 758
;; TYPE: PRT
;; ORGANISM: M. musculus
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: (1)...(758)
;; OTHER INFORMATION: Xaa - Any Amino Acid
;; S-10-117-846-16

Query Match 36.4%; Score 40; DB 9; Length 758;
Best Local Similarity 66.7%; Pred. No. 74;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 11 YTDQIERPL 19
DB 570 YSDQVORPL 578

US-09-815-915-8
;; Sequence 8, Application US/09815915
;; Patent No. US2002025931A1
;; GENERAL INFORMATION:
;; APPLICANT: Millennium Pharmaceuticals, Inc.
;; APPLICANT: Meyers, Rachel
;; TITLE OF INVENTION: 3714, 16742, 23546, AND 13887 NOVEL
;; FILE REFERENCE: 38155-20006.00
;; CURRENT APPLICATION NUMBER: US/09/815,915
;; PRIOR FILING DATE: 2001-03-23
;; PRIOR APPLICATION NUMBER: US 60/191,846
;; PRIOR FILING DATE: 2000-03-24
;; NUMBER OF SEQ ID NOS: 19

;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 8
;; LENGTH: 1244
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; S-09-815-915-8

Query Match 36.4%; Score 40; DB 10; Length 1244;
Best Local Similarity 46.2%; Pred. No. 1.3e+02;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 4 PNKGTQYTDQIE 16
DB 626 PTASEQYTDRL 638

US-09-864-761-40422
;; Sequence 40422, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;; FILE REFERENCE: Aemica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; PRIOR FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 40422
;; LENGTH: 40
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:

```

; OTHER INFORMATION: MAP TO AC015638.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.4
US-09-864-761-40422

```

```

Query Match      35.5%; Score 39; DB 10; Length 40;
Best Local Similarity 43.8%; Pred. No. 3.8;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 5 NKGTQYTDQIERPLM 20
Db 15 SGTQYTDQIERPLM 30

```

```

RESULT 14
US-09-347-331-10
; Sequence 10; Application US/09347331
; Patent No. US20010010909A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Miao, Gou Hau
; APPLICANT: Odell, Joan
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: Chromatin Associated Proteins
; FILE REFERENCE: BB-1118-C
; CURRENT APPLICATION NUMBER: US/09/347,331
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: 06/092,841
; EARLIER FILING DATE: July 14, 1998
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Gallus gallus
US-09-347-331-10

```

```

Query Match      35.5%; Score 39; DB 10; Length 428;
Best Local Similarity 36.8%; Pred. No. 57;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 1 TYPNKGTYTDQIERPL 19
Db 344 TRIENONSROYLDQIROTI 362

```

```

RESULT 15
US-09-815-242-13601
; Sequence 13601; Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848

```

```

; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13601
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13601

```

```

Query Match      35.0%; Score 38.5; DB 10; Length 232;
Best Local Similarity 52.6%; Pred. No. 34;
Matches 10; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

```

```

QY 3 PPNKGTYTDQIERPLM 20
Db 22 FLEKGYGVGDQGERPLL 40

```

Search completed: December 11, 2002, 15:28:03
Job time : 11 secs

THIS PAGE BLANK (USPTO)